DEVELOPMENT OF NEW TOOLBOXES FOR RAPID DETECTION OF SUBTLE DNA ALTERATIONS AND FOR VISUALIZATION OF RNA DYNAMICS

CHEN JIANBIN

(B. Sci. Sun Yat-Sen University)

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF BIOLOGICAL SCIENCES

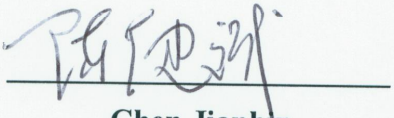
NATIONAL UNIVERSITY OF SINGAPORE

2015

THESIS DECLARATION

I hereby declare that this thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.



Chen Jianbin

12 Aug 2015

ACKNOWLEDGEMENTS

I would like to express my sincerest gratitude to my supervisor A/P Hong Yunhan, for providing an excellent study and research environment during my four-year journey of following curiosity and unveiling life's mystery. He supervised me with his resourceful ideas, ample knowledge, and proficient experiment techniques. He also encouraged me to attend international conferences and to network with talented researchers worldwide.

I am much obliged to my QE committee and my thesis advisory committee: A/P Chan Woon Khiong, A/P Yue Genhua and Prof. Wang Shu for their insightful comments and advices on my projects.

My sincere thanks go to Dr. Guan Guijun and Dr. Wang Tiansu, who provided me the knockout fish sample for my project, and to Zhang Xi for her collaboration in experiments and manuscript preparation for the PAGE-HMA project. I would also like to thank Dr. Wu Jikui, for his cooperation and guidance in experimental design and performance, data analysis, and manuscript preparation in the RNA visualization project. I thank Dr. Yuan Chenghui for his help in the LC-MS assay. My thanks also go to my labmates, Feng, Zhendong, Ni, Yongming, Mingyou, Qizhi, Fan, Narayani, Manali for their support and the wonderful atmosphere. I would like to extend my thanks to my dearest friends, Jiawei, Qianwang and Mingliang for friendship and encouragement. I also owe a debt of gratitude to Deng Jiaorong and Zeng Qinghua for medaka maintenance, Foong Choy Mei and Veronica Wong for lab administration. I am grateful to the National University of Singapore and Department of Biological Sciences for the scholarship and the support.

Finally, my special thanks go to my family for their unconditional love, especially to my beloved life partner, Siting, for her consistent support and company.

LIST OF PUBLICATIONS AND MANUSCRIPTS

Major-authored publications and manuscripts

1. **Jianbin Chen***, Xi Zhang*, Tiansu Wang, Zhendong Li, Guijun Guan, Yunhan Hong, Efficient detection, quantification and enrichment of subtle allelic alterations. *DNA Research*. 2012 Oct;19(5):423-33.
2. **Jianbin Chen***, Jikui Wu*, Hong Yunhan, Morpholino molecular beacon as a safe, stable, and specific fluorescent probe for RNA imaging *in vivo*. *Chemical Communications*. 2016, DOI: 10.1039/c5cc07124k

*, co-first author and equal contribution

Co-authored publication

1. Xi Zhang, Guijun Guan, **Jianbin Chen**, Kiyoshi Naruse, Yunhan Hong, Parameters and efficiency of direct gene disruption by zinc finger nucleases in medaka embryos. *Marine Biotechnology*. 2014 Apr;16(2):125-34.

TABLE OF CONTENTS

SUMMARY

In developmental biology research, methods for rapid detection of subtle DNA alterations generated by gene editing experiments and for non-invasive RNA visualization *in vivo* are in high demand but are lacking. This thesis aimed to develop new techniques for identifying subtle DNA alterations generated by gene editing, and for visualizing RNAs in spatiotemporal dynamic processes of developing embryos.

For the first aim, a cost-efficient procedure was established on the basis of heteroduplex mobility assay (HMA) using simple native polyacrylamide gel electrophoresis (PAGE), which allows for efficient detection, quantification and enrichment of subtle DNA alterations. This method was termed PAGE-HMA for convenient. The PAGE-HMA method was first examined in a biallelic model, where it identified the allelic alteration as distinct heteroduplex bands at high sensitivity of ~0.4%, 16-times higher than the conventional T7 endonuclease I digestion and agarose gel electrophoresis method. Then, in a multi-allelic model, PAGE-HMA could discriminate various alleles with addition or deletions of 1 - 18 bp as various heteroduplex bands, which were quantifiable by densitometry. Moreover, PAGE-HMA allows recovery and enrichment of the allele alterations via gel band recovery followed by PCR amplification, which significantly lowers the workload for cloning and sequencing the alterations. Finally, PAGE-HMA was applied to various gene editing experiments, verifying its efficiency in detecting subtle DNA alterations. This method has been verified and used by other groups for gene targeting experiments or species identification.

As for the second aim, a new type of molecular beacon was designed with the

backbone of the widely used morpholino oligonucleotides which meets the essential requirements to visualize specific RNA targets *in vivo*. In the characterization assays *in vitro*, the morpholino molecular beacon (MOMB) exhibits the insensitivity to salt concentration, the resistance to nucleases and DNA binding proteins, and the high specificity of discriminating sequence with 2 bp mismatch. Further characterization assays *in vivo*, the MOMB showed negligible toxicity, stability and specificity to RNA targets in living embryos of the fish medaka. These features make MOMB a prime candidate for imaging dynamic processes of spatiotemporal RNA expression and distribution profiles as well as dynamic processes of target RNA-containing cells such as cell migration during embryonic development and cancer cell metastasis.

LIST OF FIGURES & TABLES

LIST OF ABBREVIATIONS

%	percent
°C	degree celsius
1×	one-time dilute
AGE	agarose gel electrophoresis
bp	base pair
Cas	CRISPR-associated
CISH	chromogenic *in situ* hybridization
CRISPR	clustered regularly interspaced short palindromic repeats
DABCYL	4 - ([4 - (dimethylamino)phenyl]azo)benzoic
DNA	deoxyribonucleic acid
dpf	day(s) post-fertilization
dph	day(s) post-hatching
EMC	enzymatic mismatch cleavage
FISH	fluorescent *in situ* hybridization
FP	fluorescent protein
g	gram
GE	gene editing
h	hour
HDR	homology-directed repair
Hm	homoduplex(es)
HMA	heteroduplex mobility assay

hpf	hour(s) post-fertilization
HRM	high resolution melting
Ht	heteroduplex(es)
indel	insertions and deletions (of genome sequences)
ISH	*in situ* hybridization
L	liter
LC-MS	liquid chromatography-mass spectrometry
MALDI	Matrix-assisted laser desorption/ionization
MB	molecular beacon
mg	milligram
min	minute
mL	milliliter
mM	millimolar
MT	mutant
NHEJ	non-homologous end-joining
nm	nanometer
nM	nanomolar
PAGE	polyacrylamide gel electrophoresis
PAGE-HMA	PAGE-based HMA
PB	phosphate buffer (1 mM Na_2HPO_4 + KH_2PO_4 for pH=7, 100 mM NaCl, 3 mM $MgCl_2$)
PCR	polymerase chain reaction
RBP	rna binding protein
RNA	ribonucleic acid
RFLP	restriction fragment length polymorphism

s	second
S/N	signal-to-noise
SNP	single nucleotide polymorphism
TAGE	T7 endonuclease digestion and agarose gel electrophoresis
TALEN	transcription activator-like effector nucleases
TOF	time of fly
UTR	untranslated region
WT	wild type or wild-type
ZFN	zinc finger nucleases
µL	microliter
µM	micromolar

CHAPTER 1 Introduction

1.1 Developmental genetics: gene editing and phenotype identification

During the last 30 years, developmental biology shows a pivotal position of biological research, for it integrates the various areas of biology, including molecular and cellular biology, genetics, morphology, and evolutionary biology, providing knowledge for not only academic research but also practical applications, such as drug discovery (Zon and Peterson 2005), assisted reproductive technology (Hoozemans et al. 2004), prenatal screening (Carlson 2013), regenerative medicine (Ingber and Levin 2007), and more.

Due to the integrative nature of developmental biology, it is essential to connect the sporadic dots of knowledge into a more comprehensive network view. In this sense, developmental genetics plays a crucial role. Developmental genetics aims at unfolding how genes control the various developmental process, which links the areas of morphology and molecular biology. Therefore, in developmental genetics, two technologies are among the most important ones: gene editing, which alters the function of a specific gene, and phenotype identification, which reveals the consequence of the gene alteration. In the following sections, model organisms used as well as the technologies for gene editing and phenotype identification will be reviewed.

1.1.1 Model organisms for developmental genetics

1.1.1.1 Common model organisms

Among over 1 million animal species, only a few were chosen by developmental biologist as so called “model organisms”. The choice of model organisms considers

the cost of maintenance, the accessibility of embryos, the ease of embryonic manipulation and observation, and their genetics. Fitting these criteria, nematodes, fruit flies, sea urchin, African clawed frog, zebrafish, chick and mouse have been selected as the most commonly used model organisms representing animals ranging from invertebrates, lower vertebrates to higher vertebrates. Each of these models has their own advantages and shortcomings. Considering their relevance to human, vertebrate models are better choices, which might share more similar developmental mechanisms with human.

1.1.1.2 Medaka as a vertebrate model

In this thesis, the fish medaka (*Oryzias latipes*) was chosen to be the model organism. Medaka is a small freshwater fish that can be found in rice fields of Japan, Korea and China. It has been extensively studied for more than a century (S. R. Porazinski, Wang, and Furutani-Seiki 2011; Shima and Mitani 2004; Temminck and Schlegel 1850; Wittbrodt, Shima, and Schartl 2002), and has proved to be an excellent model organism. First, medaka is small fish that can be easily kept in large amount at low costs. Second, upon adulthood, each female medaka fish can spawn as many as 50 embryos every day, which is a significant advantage over other model organisms. Third, the large (diameter of 1 mm) and transparent embryos allow easy observations and embryological manipulations throughout the whole embryogenesis process from fertilization to hatching. Furthermore, development process of medaka embryos can be reversibly arrested at 4 ºC, which facilitats observation as well as manipulation (Sampetrean et al. 2009). Fourth, diverse wild populations of medaka with relatively high degree of polymorphism are available for genetic mapping. Fifth, medaka is highly tolerant to inbreeding, resulting in predominantly homozygous inbred lines of more than 100 generations (Kirchmaier et al. 2015).

Attracted by the convenience and versatility of medaka, researchers have been exploiting medaka comprehensively, and have built a solid platform for further research. Several embryonic stem cell lines including haploid stem cell lines of medaka have been established, which offers a unique platform for *in vitro* gene function study (Y. Hong, Winkler, and Schartl 1996; Y. Hong, Winkler, and Schartl 1996; Yi, Hong, and Hong 2009). Germ cell biology research was also conducted using medaka (Yunhan Hong et al. 2004; M. Li et al. 2009). Transgenesis is also well established for medaka including technique for transgenic line creation (Kirchmaier et al. 2013), gene and enhancer trapping (Wittbrodt, Shima, and Schartl 2002), large-scale mutagenesis (Furutani-Seiki et al. 2004), and chimera formation (Yunhan Hong, Winkler, and Schartl 1998). Moreover, whole genome sequence of medaka is available with estimated size of about 700 Mb, only a half of the zebrafish genome size (Kasahara et al. 2007).

1.1.2 Strategies and techniques for gene editing and phenotype identification

1.1.2.1 Gene editing: precise mutagenesis

Gene editing (GE) is a recent advance that enables precise customized targeted knockout of specific genes using artificial sequence specific endonucleases. These endonucleases contain both DNA binding domains and endonuclease domains. Upon the binding to target genome sequence, these endonucleases can introduce double strand breaks (DSBs) at the targeting site. Such DSBs are repaired by the targeted cells through homology-directed repair (HDR) (Jasin 1996) or non-homologous end-joining (NHEJ) (Moore and Haber 1996; Rouet, Smih, and Jasin 1994) machineries.

HDR is adopted when homologous DNA of the targeted site is presented, and the

targeted site would be repaired according to the homologous sequence (Fig. 1-1). Thus, HDR process can be used to introduce certain nucleotide substitutions or insertions to a specific site of the genome by providing an artificial "donor template" in which the sequence to be inserted or substituted are flanked by two homologous arms.

When no homologous sequence is available for HDR, NHEJ pathway will repair the DSB in a less precise manner which would result in random insertions or deletions (indels) (Fig. 1-1). Although the detail mechanisms of NHEJ is poorly understood (Deriano and Roth 2013), the imprecise repair generated by NHEJ, which could result in frame shift mutation, has been widely used for targeted gene knockout.

For GE experiments in developmental biology utilizing NHEJ mechanism, it remains a challenge to identify the individuals that carry the desired mutation. The indels introduced by NHEJ are usually too subtle to be detected. These indels are usually found to be less than 20 bp (Kim, Kweon, and Kim 2013) in a small portion of the cells. Various approaches are proposed to solve this problem, but their usage is hindered by the requirement of specific instruments or by their low sensitivity.

1.1.2.2 Phenotype identification: morphology and gene expression analysis

Upon the successful mutagenesis by GE experiment, phenotypic changes of the knockout individuals are the key to understand the functions of the knockout genes. To identify these phenotypic changes, morphological and gene expression profile analysis are of the focus.

1.1.2.2.1 Techniques for morphology analysis

To unravel the morphological changes after GE experiments, cell-labeling methods could be used for cell fate mapping, lineage analysis and clonal analysis (Slack,

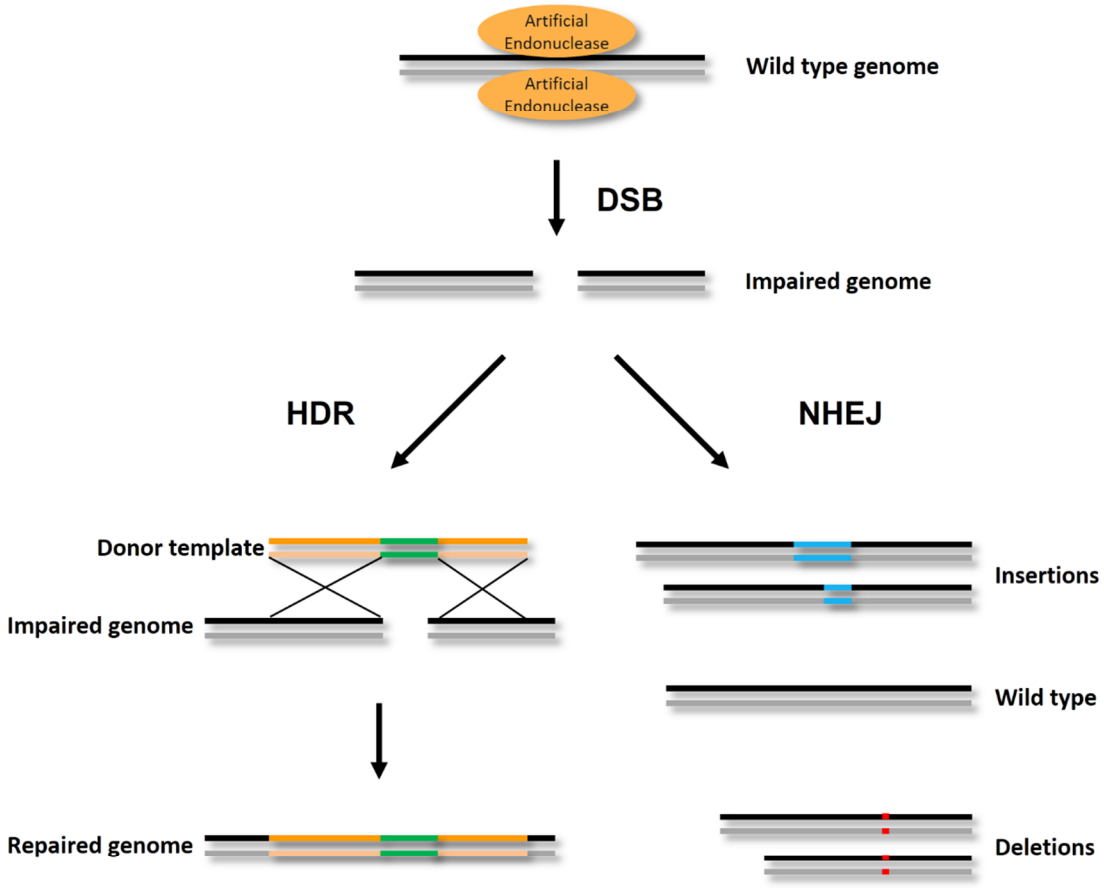


Figure 1-1. HDR and NHEJ in gene editing. The DSBs introduced by sequence-specific artificial endonucleases are repaired by HDR or NHEJ. With the presence of donor templates which contain homologous arms of the target region, HDR can introduce nucleotide substitutions or insertions as desired. If no donor templates are provided, NHEJ would repair the DSB by directly joining the ends resulted from DSB, which could produce random insertions or deletions to the targeted site.

2009). One of the most commonly used cell-labeling tools is fluorescent dextran. As highly water-soluble polymers of glucose, dextran molecules can be produced in different sizes from 1 to 2000 kDa, and can be modified with a wide range of fluorescent dyes or other chemical units, thus providing a large variety of cell-labeling tracers. They have been applied to fate mapping, lineage analysis and endocytosis experiments in different species (Doherty and McMahon, 2009; Stehno-Bittel et al., 1995; Strehlow and Gilbert, 1993; Thorball, 1981; 2009). It has been well established that these dextran tracers are metabolically inert, nontoxic to cells or organisms, and uncharged (Thorball 1981). In the absent of endocytosis, dextran molecules of more than 40 kDa are not permeable to living cell membranes or organelle membranes (Stehno-Bittel et al., 1995). In this thesis, Texas Red dextran was used as a non-toxic indicator for microinjection. A major weakness of the fluorescent dextran is that it has no specificity on cell types, i.e they show fluorescent signal regardless of the cell types.

To achieve cell type specific cell-labeling, genetic labeling techniques are used. There are two variants of genetic labels, both require establishment of transgenic lines with reporter genes, such as fluorescent proteins, β-galactosidase, and alkaline phosphatase. The first approach utilizes transgenic lines with universal expression of reporter genes. Such transgenic lines can serve as a donor in chimera forming experiments, in which all progenies from the donor can be identified live or *in situ* (Yunhan Hong, Winkler, and Schartl 1998; Rembold et al. 2006). The second approach fused cell-type specific promoter and 3' UTR (untranslated region) with the reporter genes, creating transgenic lines that express the reporter protein in specific type of cells (Tanaka et al. 2001). Genetic labeling approach enables the stable and unambiguous tracking of a certain type of cells. However, the

establishment of transgenic animal lines is laborious and time-consuming. Moreover, since cryopreservation of transgenic lines is not applicable to all model organisms, it costs money and time for proper maintenance of these lines.

1.1.2.2.2 Techniques for gene expression analysis

Various approaches for detecting gene expression level are established on both mRNA and protein level. For the temporal expression profiles of genes, qPCR and western blotting are the two primary static methods used to investigate mRNA and protein expression level, respectively. Both methods require extraction of RNA or protein from the sample, consuming the sample in each extraction. Moreover, to gain a complete expression profile of the whole development procedure, multiple samples from different time point are to be prepared. For the spatial expression profile, *in situ* hybridization of RNA and immunohistochemistry are the two most commonly used procedure. These procedures can be performed on the basis of sections or whole-mount embryos. They can reveal the spatial expression profile of certain genes at a specific time point, yet they could not provide real-time information of the profile change. In addition, the available antibodies for the genes of interest are limited or difficult to prepare, which hinder their application.

To gain a real-time spatiotemporal expression profile, transgenic lines are established with fluorescent protein sequence fused to the gene specific promoter and UTR regions as described in 1.1.2.2. Such method can only reveal the expression profile on protein level.

To date, no technique has been reported to visualize RNA expression in real-time. However, expression profile on RNA level also serves an important role in developmental genetics. For example, maternal RNAs control the early stages of the development (Gavis and Lehmann 1992; Rebagliati et al. 1985; Yoon,

Kawakami, and Hopkins 1997), and their spatiotemporal distribution could be important clues of their functions. For another example, miRNAs have shown to be essential for embryonic development (Alvarez-Garcia and Miska 2005; Kloosterman et al. 2006; Tay et al. 2008), yet no existing approach is available for real-time visualization of miRNA in living embryos.

1.1.3 Challenges in gene editing and phenotype identification

In sum, three challenges in developmental genetics were introduced above. First, a rapid approach is lacking for the detection of indels from GE experiment subjects. Second, easy and efficient tools are demanded to track morphological change of certain types of cells in developing embryos. Third, the technique for real-time RNA visualization in living embryos is desired but lacking. Actually, if a real-time RNA visualization technique is available, the second challenge could be solved by monitoring the expression of a cell-type-specific marker gene. Therefore, in the following sections, current strategies for indel identification and for RNA visualization will be introduced in detail, and new approaches for efficient solutions will be proposed.

1.2 Genome editing in developmental biology

For efficient GE experiments, various tools for generating DSB have been developed that are based on the engineered endonucleases, including Zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and CRISPR-Cas. Meanwhile, several approaches are available for detecting indels generated by the GE experiments. In the following sections, these engineered endonucleases and current approaches for indel detection will be introduced.

1.2.1 Engineered endonucleases

1.2.1.1 ZFN

Zinc finger nucleases (ZFNs) are the first widely used artificial DNA nucleases for gene editing. They consist of several zinc finger DNA-binding domains engineered for recognition of the target sequence (Beerli and Barbas 2002) and a FokI endonuclease fused to the recognition domain for cleavage at the target site (Porteus and Baltimore 2003). The FokI is only functional when forming dimers, so a pair of ZFNs are used in a GE experiment. The DNA-binding domain, Cys2-His2 zinc finger domain, is among the most common DNA-binding motifs in eukaryotes. Each domain contains about 30 amino acids with a zinc ion, forming a ββα configuration which can bind to the major groove of a specific 3 bp DNA. Variation of this domain allows recognition of various DNA triplets. By combining 6 zinc finger domains according to the site of interest, a specific 18 bp sequence can be targeted. To this end, preselected libraries of the zinc finger domains are established covering nearly all 64 possible DNA triplets, and commercialized ZFN systems are available for customized gene editing purposes. The ZFN approach was the sole customizable site-specific gene editing strategy for many years. It is proved to be adaptable to mediate gene knockout in diverse organisms including fruit fly, zebrafish, mouse, and human (Carroll 2011; Gaj, Gersbach, and Barbas 2013). In addition, ZFNs facilitate HDR process after the introduction of DNA cleavage, resulting in 10-fold increase of the integration of HR cassette (Bibikova et al. 2001). Though ZFNs enable exciting possibilities for gene knockout analysis, construction of functional ZFNs has proven to be difficult due to the context-dependent nature of individual ZFN units (Wolfe, Nekludova, and Pabo 2000), which could prevent their wider applications.

1.2.1.2 TALEN

As an alternative genome editing tool, transcription activator-like effector nucleases (TALENs) have been studied and applied widely in recent years. TALENs emerged from the TALE proteins discovered in plant pathogenic bacteria *Xanthomonas* (Boch et al. 2009). Similar to ZFNs, TALENs are artificial fusion proteins with DNA binding domains and FokI endonuclease domains, and are works in pairs. Unlike ZFN, the DNA binding domains of TALENs consist of tandem repeats of 33-35 amino acids each, and each unit of repeats recognizes one nucleotide, specified by the 12^{th} and 13^{th} amino acids (NI, adenine; HD, cytosine; NG, Thymine; NN, guanine and adenine) (Moscou and Bogdanove 2009). These units have fewer context-dependent effect (Reyon et al. 2012). Hence, drawing upon this simple one-to-one recognition principle and the modular nature of TALE proteins, customized TALENs can be easily assembled (Doyle et al. 2012; Miller et al. 2011; Reyon et al. 2012). TALENs have been successfully adopted for GE experiments in various organisms (Joung and Sander 2012; T. Wang and Hong 2014).

1.2.1.3 CRISPR-Cas

Recently, a new system of CRISPR-Cas has been established providing more flexibility and robustness to GE. CRISPR is short for clustered regularly interspaced short palindromic repeats, while Cas for CRISPR-associated. CRISPR systems were identified in many bacteria as adaptable immune systems which degrade specific foreign RNA or DNA (Fineran and Charpentier 2012; Wiedenheft, Sternberg, and Doudna 2012). Type II CRISPR system can introduce DSB to specific DNA target via a single Cas9 endonuclease, thus it is used for the development of engineered CRISPR-Cas GE system. Different from protein-DNA interaction based ZFNs and TALENs, Cas9 endonuclease relies on a RNA complex of CRISPR RNA (crRNA)

and transactivating crRNA (tracrRNA) to perform cleavage. For research application, such RNA complex has been simplified into a single guide-RNA (gRNA) molecule (H. Yang et al. 2013).

1.2.2 Detection of genome editing events

One of the major challenge in GE experiments is to efficiently detect and identify indels introduced by the engineered endonucleases. Most indels from GE experiment are found to be less than 20 bp (Kim, Kweon, and Kim 2013), which are hardly detectable using standard agarose gel electrophoresis (AGE). Moreover, in a typical GE experiment, only a small percentage of the cells carry the desired indels, which would be difficult to identify using Sanger sequencing. For GE in cell lines, a HDR donor plasmid with reporter genes or antibiotic resistant genes is usually included to enrich cells with successful gene targeting events. Whereas for GE in embryos intended for gene knockout line establishment, new strategies are demanded since the small indels with a relative low abundance in the embryo are not likely to produce identifiable phenotypic change. In the following sections, commonly used methods will be reviewed.

1.2.2.1 Enzyme mismatch cleavage

One of the simple and widely used method is enzymatic mismatch cleavage (EMC) followed by electrophoresis (Babon, Youil, and Cotton 1995; Youil, Kemper, and Cotton 1995). This method depends on a certain kind of enzymes which identify mismatches of double stranded DNA and create DSB at the mismatch sites. T7 endonuclease I (T7EI) is a typical member of these enzymes (Vouillot, Thélie, and Pollet 2015). Using T7EI followed by AGE is a routine protocol for indel identification. This protocol begins with amplifying the targeting region by PCR from candidate samples as well as from a WT sample. Then the candidate amplicons are

mixed with amplicons from the WT sample followed by a heteroduplex formation procedure. The products are then subject to digestion by T7EI and AGE analysis. For indel containing samples, more than one band would be seen on AGE, while for the WT sample, only one band would be seen. Other enzymes might be used such as CEL (Oleykowski et al. 1998; Qiu et al. 2004) and ENDO (Triques et al. 2008). EMC is very cost-effective and it requires no specific instrument. Nevertheless, this method is limited by its low sensitivity even when using capillary electrophoresis, which could only detect mutations of more than 5% in the PCR product pool (Chen et al. 2012; Vouillot, Thélie, and Pollet 2015).

1.2.2.2 Restriction fragment length polymorphism

Restriction fragment length polymorphism (RFLP) is another commonly used approach for identifying allelic mutations (Pourzand and Cerutti 1993; Chakravarti et al. 1998). This approach can detect mutations located in a restriction enzyme recognition sequence through the resulting resistance to the cleavage by the respective restriction enzyme. Hence, in GE experiments using this approach, the cleavage site of engineered endonucleases must include certain restriction enzyme recognition sequences. To perform RFLP analysis, PCR amplicons of the targeting region are prepared and digested by restriction enzymes accordingly. Digestion products are then analyzed by AGE, and the undigested samples can be sequenced to identify the indels. RFLP is also cost-effective and has been widely used for screening indels from GE experiments (Feng et al. 2013; Hruscha et al. 2013, 9; H. Yang et al. 2013). The major shortcoming of RFLP is that it limits the selection of targeting region for GE experiment.

1.2.2.3 High-resolution melting

High resolution melting (HRM) analysis is an alternative but less popular approach

for detecting mutations (Dahlem et al. 2012; Parant et al. 2009). Upon the coexistence of WT and mutant alleles, a denaturing and annealing process would produce specific heteroduplexes (Ht) from the original homoduplexes (Hm). Hence, mutations can be detected via identifying Ht. HRM approach is capable of discriminating Ht and Hm by their thermostability difference. In HRM, short PCR amplicons (90-120 bp) of candidate samples along with a WT sample are to be prepared. Unlike the EMC and RFLP, HRM does not require enzymatic reaction after PCR. The amplicons from candidate samples are mixed with that from WT sample, followed directly by the analysis of their thermostability during a denaturing and rapid annealing process. This labor-saving approach allows for the high-throughput screening in large-scale GE experiments (Thomas et al. 2014). Nevertheless, specific instruments are necessary for HRM which prevents its wider use.

1.2.2.4 Heteroduplex mobility assay

Heteroduplex mobility assay (HMA), was used widely to detect mutations in disease genes (Glavac and Dean 1995; Highsmith et al. 1999) or to discriminate different strains of bacteria or virus (Kostrikis et al. 1995; Leys et al. 1999; Zou 1997). HMA can separate Ht and Hm emerged from the denaturing and annealing of WT and mutant alleles by their mobility differences on polyacrylamide gel electrophoresis (PAGE) or other electrophoreses. It does not require special instrument and can be easily adopted by common biology labs. Before my study (Chen et al. 2012), there is no report on using HMA for the detection of indels from GE experiment. In this thesis, the PAGE-based HMA method (PAGE-HMA) will be examined for rapid detecting, quantifying and enriching indels generated by GE experiments. And a new GE work flow based on this PAGE-HMA method will be developed.

1.3 Visualization of RNA

As discussed in section 1.1.2.3, real-time RNA visualization could provide valuable information of gene and miRNA expression profiles during development processes, and could also be used to achieve cell type specific labeling. However, efficient approach for real-time visualization of RNA *in vivo* is lacking. The conventional strategy used for visualization RNA is *in situ* hybridization (ISH), which can only gain a static profile for each sample, unless multiple ISH are carried out with samples from different developmental stages. In recent years, numerous strategies have been developed to visualize RNA dynamics in living cells, using RNA binding proteins (RBP), RNA aptamers and molecular beacons (MBs). In developing embryos, although fluorescent proteins can be used to trace the expression of specific proteins, the visualization of specific RNA remains a big challenge. In the following sections, the strength and weakness of these major strategies of RNA visualization will be reviewed.

1.3.1 Static visualization of RNA by *in situ* hybridization

Developed more than 40 years ago (Buongiorno-Nardelli and Amaldi 1970; Gall and Pardue 1969; John, Birnstiel, and Jones 1969), ISH visualizes a specific segment of RNA via a complementary strand of nucleic acid or oligonucleotide to which a reporter molecule is attached. Radioactive isotopes were first used as the reporter molecules, and were replaced later by hapten labels such as biotin and digoxigenin, or by fluorescent labels. According to the reporter molecule used, ISH can be classified as two types: chromogenic *in situ* hybridization (CISH) that use hapten labels (Tanner et al. 2000; Zhao et al. 2002), and fluorescent *in situ* hybridization (FISH) that used fluorescent labels (Vize, McCoy, and Zhou 2009; Wiegant et al. 1993).

1.3.1.1 CISH

In CISH, hapten labels are used as reporter molecules which require immunohistochemistry chromogenic process to exhibit visible signals. The whole CISH procedure begins with the synthesis of hapten attached probes and the preparation of the sample (fixation and permeabilization). Then it is followed by the binding of probes to target mRNA and the washing of excessive unbound probes. Next, antibody-bound chromogenic enzymes are added to bind the haptens. Finally, the chromogenic substrates are added for a certain time to develop the signal. This process can be applied to section samples on glass slides, or to whole embryos samples in the procedure termed whole mount *in situ* hybridization (WISH). The major advantage of CISH is its ability to visualize RNAs of low concentration due to its amplification effect. It is cost-efficient regarding the reagents used and the normal bright-field microscopes for imaging. Reagents for CISH are usually stable and samples after CISH can be stored for a long time before reexamination. Nevertheless, CISH is a less precise method regarding its spatial resolution. It is incapable of subcellular localization of RNA since the chromogenic staining could cover a large area around the target RNA. Also, in WISH, the chromogenic staining of the outer layer of the embryo could mask the information from the inner layers. Moreover, it is a labor-intensive and error-prone procedure which could take up to weeks to optimize.

1.3.1.2 FISH

In FISH, two approaches are used: indirect labeling and direct labeling. For indirect labeling, probes with hapten labels are synthesized first, akin to that for CISH. Then, fluorophore-attached antibody to the corresponding hapten is used to amplify and exhibit fluorescent signals. For direct labeling, fluorophores are linked directly to the probe, thus no antibody is required (Wiegant et al. 1993). One major advantage of

FISH approach is the high resolution and signal to noise ratio which allow subcellular localization of RNA (Swiger and Tucker 1996). Another advantage of FISH is that it allows the detection of multiple target within one sample (Vize, McCoy, and Zhou 2009).

1.3.2 Dynamic visualization of RNA

To gain a more comprehensive understand of the elaborate development processes, static visualization approaches are not enough. Technologies are demanded for real time visualization of RNA in living organisms so as to elucidate the molecular basis of the highly dynamic gene expression and morphological change of the embryos. To date, although no approach is reported for the visualization of RNA dynamics in developing embryos due to their innate complexity, various approaches are available for depicting the spatiotemporal gene expression profile in living cells. In the following sections, representative strategies for dynamic visualization of RNA in living cells will be introduced.

1.3.2.1 RNA binding protein (RBP)

1.3.2.1.1 RNA binding proteins tagged with fluorescent protein

The first strategy for tagging RNA utilizes RNA-binding proteins (RBPs) tagged with Fluorescent proteins (FPs) (Fig. 1-2 A). RBPs chosen for this purpose can bind to certain RNA motifs with high specificity and affinity, while neither the RBPs nor the RNA motifs should be in the targeted cells (Sanjay Tyagi 2009; Urbanek et al. 2014). FPs are arguably the most widely used biocompatible fluorophores that can be easily fused to almost any protein of interest. The characteristics and applications of FPs in developmental biology can be found in previous reviews (Dean and Palmer 2014; Pantazis and Supatto 2014; Yu et al. 2003).

Two different approaches were reported in FP tagged RBP for RNA visualization.

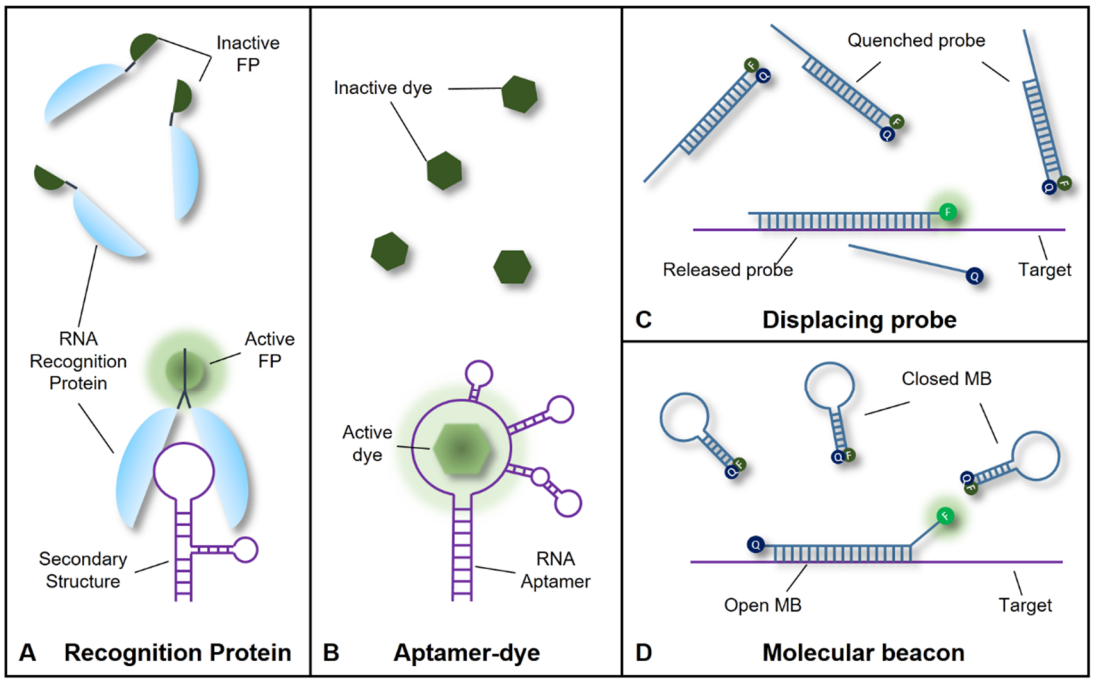


Figure 1-2. Principles of popular RNA detection systems. (A) RNA recognition protein. (B) Aptamer system. (C) Displacing probe. (D) Molecular beacon. Purple lines denotes target RNA. FP, fluorescent protein; Q, quencher; F, fluorophore; MB, molecular beacon.

The first one use the whole intact FPs (Bao, Rhee, and Tsourkas 2009; Bertrand et al. 1998; Sanjay Tyagi 2009). In this approach, multiple binding sites (>10) are required for one single RNA molecule in order to achieve a high local concentration of FPs for a recognizable signal over the background signal of the unbound FPs. MS2 system is the representative system of this approach.

The second improved approach uses split FPs, which are two complementary fragments of the original FP that fused to two RNA-binding proteins (Bao, Rhee, and Tsourkas 2009; Sanjay Tyagi 2009; Valencia-Burton et al. 2007). These split FPs are non-fluorescent until their respective RNA-binding proteins bind to adjacent sequence, where the two fragments can assemble and restore the fluorescence (B. Wu, Chen, and Singer 2014).

1.3.2.1.2 MS2 system

As the first FP tagged RBP system, MS2 system utilizes the MS2 coat protein from R17 or MS2 bacteriophage (Bertrand et al. 1998; Fusco et al. 2003). The MS2 coat protein specifically binds to a binding site of 19-nt RNA stem-loop structure originated from the bacteriophage's RNA operator (Bertrand et al. 1998, 1) and form a stable dimer at the recognition site. To visualize the mRNA of a specific gene, multiple binding sites of MS2 are inserted into the mRNA sequence. For example, 24-48 MS2 binding sites were inserted into the 3' UTR region for visualization of single endogenous mRNA molecule during the export of mRNA from the nucleus (Grünwald and Singer 2010), in living neurons (Lionnet et al. 2011), or in living mouse (Park et al. 2014). Other RBP systems have also been reported for RNA dynamic visualization using FP tagged RBPs, such as the PP7 (Larson et al. 2011) and λN22 (Daigle and Ellenberg 2007) system.

1.3.2.1.3 Split protein systems

The major drawback of the intact FP tagged system is the high background from unbound FPs. To solve this problem, split FPs approach was developed and has proved to significantly improve the signal-to-noise (S/N) ratio (Bao, Rhee, and Tsourkas 2009). The first report of such system used MS2 coat protein together with another RBP (Rackham and Brown 2004). Later, another RBP called eIF4A was reported to contain two domains, each of which show strong affinity to one side of the target RNA motif (Valencia-Burton et al. 2007). Thus it is naturally suitable for the split FP approach. In addition, split FPs with three parts are also developed to further minimize the background fluorescence from self-assembly of the split parts (Cabantous et al. 2013).

1.3.2.1.4 Limitations of the RBP strategy

The major limitation of the RBP strategy is that it entails the insertion of the multiple RBP binding sites, which requires the invasive engineering of the target gene. This means the laborious process of creating a transgenic line for developmental studies. In addition, the inserted sequences might alter the natural behavior of the original mRNA and result in RNA clumps (Itzkovitz and van Oudenaarden 2011). Moreover, it is not applicable for the detection of the short RNA such as miRNA or snRNA.

1.3.2.2 Aptamer systems

1.3.2.2.1 Principles of aptamer systems

The second system does not rely on the use of RBPs, but uses certain small organic fluorophores for recognition and visualization of RNA (Fig. 1-2 B). These fluorophores contain constituents that can rotate or vibrate freely. Consequently, they show negligible fluorescence in dissociative condition because they dissipate the excitation energy to intramolecular motions (Babendure, Adams, and Tsien 2003;

Sanjay Tyagi 2009). To restore their fluorescence, restrictions on these constituents are required. RNA aptamers, which are artificial structural RNA motifs originating from *in vitro* selection experiments, can specifically bind to these fluorophores and restrain their intramolecular motion. Upon binding with the corresponding aptamer, fluorescence of these fluorophores can be restored with more than 2000-fold increase (Babendure, Adams, and Tsien 2003; Hermann and Patel 2000). Hence, with the insertion of aptamer sequences, the target RNA can be visualize by these fluorophores in living cells (Dolgosheina et al. 2014; Sato et al. 2015; Song et al. 2014). Among these aptamer-fluorophore system, spinach is one of the most famous.

1.3.2.2.2 Spinach system

The Spinach system was first reported as a mimic of GFP for RNA visualization (Paige, Wu, and Jaffrey 2011). The fluorophores for Spinach system are derivatives of 4-hydroxybenzlideneimidazolinone (HBI), which is the fluorophore in GFP. RNA aptamers for each derivatives were produced to mimic the GFP that stabilizes and activates the fluorescence of HBI. The derivatives in the Spinach system provide a wide range of fluorescence wavelengths, enabling multi-color labeling (Paige, Wu, and Jaffrey 2011; Song et al. 2014).

1.3.2.2.3 Limitations of the aptamer strategy

Akin to strategies using RBPs, visualizing RNA with aptamer-dye systems demands the insertion of aptamer to the gene of interest. Moreover, for a complete restoration of fluorescence, a strictly correct conformation of the aptamer is required, which might be difficult to maintain in the complex environment of cytoplasm, resulting in low or no fluorescence.

1.3.2.3 Hybridization probes

For non-invasive strategies of RNA visualization, hybridization probes have been

designed which requires no engineering of genes and enable detection of short RNA in living cells. Generally, these probes are oligonucleotides complementary to the target sequence, and are attached with fluorophore and quencher pairs that show fluorescence only upon correct binding with their target sequence. Hybridization probes can be divided into two groups: displacing probes with two complementary strand, and molecular beacons with one stem-loop strand.

1.3.2.3.1 Displacing probes

Displacing probes, which is also call Yin-Yang probes (Q. Li et al. 2002), are composed of two complementary oligonucleotides of different lengths, with a fluorophore on one strand and a quencher of the fluorophore on the other strand (Fig. 1-2 C). The longer target-complementary strand usually carries the fluorophore. Without binding to the target, the probe exists as a double-stranded dimer in which the fluorophore is quenched due to the close proximity of the fluorophore and the quencher. In the presence of the target, the shorter quencher strand will be displaced by the target sequence, thus releasing the fluorophore from the quencher. Displacing probes has been used for real-time PCR genotyping (Cheng, Zhang, and Li 2004) as well as mRNA detection in living cells (Seferos et al. 2007). Displacing probe approach also allows multiple RNA visualization when using different fluorophore-quencher pairs for different targets. However, displacing probes suffer from a lower specificity comparing to other RNA visualization approaches. In addition, displacing probes require the additional hybridization and purification procedure to produce probe duplexes prior to their application, which increases the difficulty in preparing functional probes.

1.3.2.3.2 Molecular beacons

Molecular beacons (MBs) are single-stranded probes with fluorophore on one end and

quencher of the fluorophore on the other (Medley and Zhu 2013; Monroy-Contreras and Vaca 2011; S. Tyagi and Kramer 1996) (Fig. 1-2 D). They contain complementary short sequences on both ends so that stem-loop structures are formed to maintain close contact of fluorophore and quencher when no target RNA exists. The complementary sequences to the target are set in the middle of the oligonucleotide. Upon the recognition and hybridization of MBs and target RNA, the stem-loop structure would open, and the fluorophore would be freed from the quencher, emitting signals of more than 200-fold compared to the quenched state (Bao, Rhee, and Tsourkas 2009). MBs can be designed with different backbones, which result in different characteristics. In the following section, detailed features and the most frequently used backbones of MBs will be covered.

1.3.3 Molecular beacons (MB): features and backbones

The design of molecular beacon was first proposed and used to monitor PCR amplicons in real time (S. Tyagi and Kramer 1996). Later, MBs were widely used for real-time PCR product detection (Sanjay Tyagi, Bratu, and Kramer 1998), SNP detection (Piatek et al. 1998) and RNA detection in living cells (Bratu et al. 2003; Sokol et al. 1998; Yeh et al. 2008). These conventional MBs use DNA as the backbone, with limitations such as the low specificity and stability in living cells. Therefore, new backbones including 2'-O Methyl RNA (2'-OMe-RNA), locked nucleic acid (LNA), and peptide nucleic acid (PNA) have been introduced to the backbone family of MB (Fig. 1-3). In the following sections, principles of MB design, backbone choice, and the challenges of MB application will be introduced.

1.3.3.1 Basic characteristics of molecular beacons

In the design of a functional MB, the thermodynamics of the sequence is crucial, which determines the balance between specificity and hybridization rate. Generally

speaking, more stable (longer length or higher C-G content) stems of MB result in better specificity but worse hybridization rate in that more stable stems maintain stronger stem-loop conformation which is harder to open by either mismatch or perfect-match sequences (Tsourkas et al. 2003). Therefore, a typical MB consists of 5-7 nucleotides as stem and 15-25 nucleotides as loop (Monroy-Contreras and Vaca 2011), which usually exhibit a Tm of 55-65 °C for the stem and Tm of 70-80 °C for the duplex of MB and target. To achieve better hybridization rate without compromising the specificity, "shared stem" strategy is adopted in which one arm of the stem also contributes to the recognition sequence following the loop (Tsourkas, Behlke, and Bao 2002a). In addition, guanine bases have certain quencher effect and should be considered in MB design (Medley and Zhu 2013; Seidel, Schulz, and Sauer 1996).

1.3.3.2 Backbones for molecular beacons

1.3.3.2.1 DNA

Conventional MBs use DNA backbone (Fig. 1-3), which is prone to be degraded in living cells with a half-life as short as 15-20 min (J. J. Li, Geyer, and Tan 2000; Uchiyama et al. 1996). MBs with DNA backbone are also affected by nucleic acid binding proteins, which could disrupt the stem-loop structure and cause high background signals (C. J. Yang et al. 2006). Moreover, the Watson-Crick hydrogen bond of the stem of DNAMB requires considerable concentration of salt to maintain, so as that of the DNAMB-target hybrids (S. Tyagi and Kramer 1996). In live-cell condition, DNAMB might suffer from low specificity due to the low salt concentration environment of the cytoplasm. In addition, native RNA in live-cell condition could form complex structures which hurdle the binding of complementary DNA. To overcome these problems, new backbones have been used in MBs.

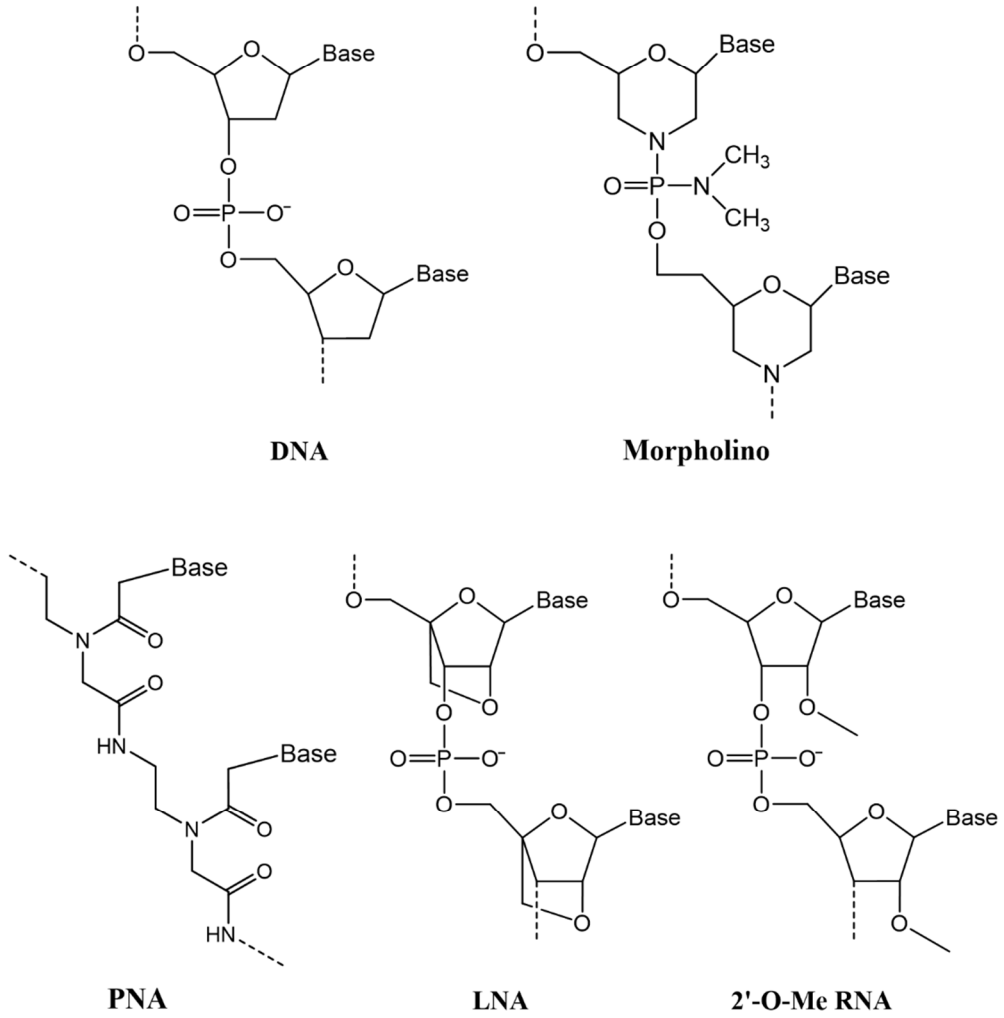


Figure 1-3. Different backbones of MB.

1.3.3.2.2 2'-OMe-RNA

2'-OMe-RNA (Fig. 1-3) is a nucleic acid analog with high affinity to RNA targets which allows binding of structured RNA region (Majlessi, Nelson, and Becker 1998). Moreover, it show high resistance to nucleases (Majlessi, Nelson, and Becker 1998). It has been used as MB backbone alone (Molenaar et al. 2001), or with several incorporations of LNA for RNA imaging in living cells (Catrina, Marras, and Bratu 2012).

Although claimed to be nuclease-resistant, its half-life, according to the disappearance of the starting oligos, was reported to be only 68min in cell culture medium with fetal serum (Morvan et al. 1993). One of my preliminary experiments in medaka embryos also indicates the low stability of 2'-OMe-RNA displacing probes *in vivo* (based on fluorescent intensity change). In addition, MBs with 2'-OMe-RNA backbone still open and show nonspecific signals in living cells (Tsourkas, Behlke, and Bao 2002b), possibly due to protein binding.

1.3.3.2.3 LNA

LNA (Fig. 1-3) is another nucleic acid analog which has remarkably high affinity to DNA and RNA (Singh et al. 1998). In addition, LNA is shown to be stable in living cells with resistance to nucleases and binding proteins (C. J. Yang et al. 2007). LNA also exhibits low toxicity to living cells (Zhang et al. 2011). Therefore, LNA has been chosen as ideal backbone for MBs in live-cell imaging (Catrina, Marras, and Bratu 2012; L. Wang et al. 2005; C. S. Wu et al. 2012).

Nevertheless, the high affinity of LNA also show certain limitations. The incorporation of a single LNA base to a duplex can substantially increase the thermo stability of the duplex. Consequently, MB with full LNA backbone show a relatively low hybridization rate (C. J. Yang et al. 2007). Hence, LNA are often used as LNA-

incorporated probes with other backbone such as 2'-OMe-RNA (Catrina, Marras, and Bratu 2012), which could result in less resistance to nucleases and binding proteins.

1.3.3.2.4 PNA

PNA (Fig. 1-3) is an oligonucleotide mimic with the N- (2-aminoethyl) glycine backbone, which is electrically-neutral. It was first introduced to MB design for a surface immobilized probe (Ortiz, Estrada, and Lizardi 1998), and was later used for short stemless MB (Kuhn et al. 2002; Socher et al. 2008). Due to its lack of negative charge, PNA has a higher affinity and rapid hybridization kinetics to target sequence than DNA (Hyrup and Nielsen 1996). PNA is also resistant to nucleases and has low toxicity in living cells (Kam et al. 2012; Karkare and Bhatnagar 2006). Moreover, the neutral backbone of PNA also enables the binding of target in low salt concentration (Nielsen and Egholm 1999).

The major drawback of PNA backbone is its low solubility in water. As the total length or purine: pyrimidine ratio increases, the solubility of the PNA oligomer decrease rapidly (Hyrup and Nielsen 1996). Thus, the PNA oligomer is limited to less than 18 nt with less than 60% of purine (Xi et al. 2003). Such short sequence for recognition could compromise the specificity of the probe.

1.3.3.3 Challenges of molecular beacons for *in vivo* application

Visualization of RNA dynamics is of vital significance in understanding the molecular basis of the delicate developmental process and have been demanded by developmental biologists. However, it is a challenging task since the developing embryo is a highly dynamic system with dramatic gene expression and morphological changes (Arbeitman et al. 2002; Tadros and Lipshitz 2009), which is unfavorable to the existing RNA probes. For example, the development process in the first several hours are controlled by maternal mRNAs and proteins, which are loaded into the egg

during oogenesis. These maternal mRNA and proteins are estimated to involve 40-75% genes of the whole genome (Aanes, Collas, and Aleström 2014; Q. T. Wang et al. 2004; Wei, Angerer, and Angerer 2006), which bring about the complexity of the intracellular environment. Nevertheless, in a very short period of time, usually less than 10 hours, more than 30% of these maternal transcripts are actively degraded via various cascades and nucleases (Ferg et al. 2007; Semotok et al. 2008; Voeltz and Steitz 1998). These cascades and nucleases might be hazardous to the protein or nucleotide based probes. In addition, during the rapid degradation of the maternal mRNAs, zygotic genome is activated and more than 12% of the genome is involved (Hamatani et al. 2004; Mathavan et al. 2005), while during the whole embryonic development, over 50% genes show their expression peaks (Aanes, Collas, and Aleström 2014; Arbeitman et al. 2002), let alone the regulatory short RNAs. With exposure to such a vast amount of RNA sequences, the specificity of the probes is challenged.

To achieve rapid and non-invasive RNA visualization, MB seems to be the best choice among the existing methods. Nevertheless, the highly dynamic environment of developing embryos raises big challenge to the existing MBs. The first challenge is the low salt concentration environment of the cytoplasm (Slack, Warner, and Warren 1973). For conventional MBs, the maintenance of the stem-loop conformation as well as the recognition of target sequence requires require certain ion strength. Though PNA could work under low salt condition, the poor solubility hinder its application. The second challenge is the instability of MBs in developing embryos. Conventional MBs with DNA backbone are sensitive to nucleases in living cells and are incapable of long-term RNA tracking. For MBs with PNA or 2'-OMe-RNA backbone, to my best knowledge, the longest survival time reported in living cells is around 2 days

(Kam et al. 2012). To achieve long-term RNA tracking during the embryonic development process, MBs with a new backbone is demanded.

1.3.4 Morpholino as a potential molecular beacon backbone

1.3.4.1 Low toxicity, high stability and specificity

Another popular gene knockdown technology utilizes antisense Morpholino oligonucleotides (MOs) (Ekker and Larson 2001). MOs are DNA analogs with a 6-member ring backbone which is resistant to degradation in living cells (Summerton and Weller 1997). Upon delivered to the cells, MOs can specifically bind to the exon-intron boundary of the pre-RNA and block its correct splicing. It can also bind to the translation initiation region and block translation. In both way, the targeted gene can be knocked down. MO was first administered to *Xenopus* embryos (Heasman, Kofron, and Wylie 2000), and later to sea urchin (Howard et al. 2001), zebrafish (Nasevicius and Ekker 2000) , medaka (M. Li et al. 2009), and chick(Kos et al. 2001), demonstrating its efficiency and biocompatibility. In this thesis, morpholino was used in a novel way for visualizing RNA *in vivo*.

Morpholino oligos (MOs) (Fig. 1-3) are commonly used for gene knockdown experiments in zebrafish and medaka. MO, which is a DNA analogs with a 6-member ring backbone, possesses several excellence traits as a potential MB backbone. First, it has a high solubility, which is more than 26g per 100g of water (Summerton and Weller 1997). Second, through a MALDI-TOF mass spectrometry (MS) based detection method, MO is proved to be resistant to a wide range of DNases, RNases, proteases, and human plasma (Hudziak et al. 1996). Third, morpholinos show high affinity to RNA targets in thermal stability examinations (Stein et al. 1997; Summerton and Weller 1997). Morpholino can also access stable RNA secondary structures (Summerton 1999). Fourth, morpholinos show good specificity to targets in

cell-free solutions as well as in living cells (Stein et al. 1997; Summerton and Weller 1997). Fifth, morpholino show no acute toxicity in a very high dose (700 mg/kg) when injected intravenously into mice (Summerton and Weller 1997). Although certain off-target effects can be seen occasionally (Bedell, Westcot, and Ekker 2011; Gerety and Wilkinson 2011), morpholino is not considered as highly toxic to embryos. In sum, morpholino oligos appear to meet the requirements for a low toxicity, high stability and specificity RNA probe for *in vivo* visualization.

1.3.4.2 Usage of morpholino in developmental biology

The demand for easy and effective reverse genetic approach in developmental biology is rising rapidly in this genomic era. As a well-established antisense technique initially for therapeutic applications, morpholino drew the attention of developmental biologists early in 2000 (Heasman, Kofron, and Wylie 2000), and have since been successfully adopted for gene knockdown experiments in a wide range of model organisms including sea urchin (Howard et al. 2001), *Xenopus* (Nutt et al. 2001), zebrafish (Nasevicius and Ekker 2000; Ross et al. 2001), chick (Kos et al. 2001), and mouse (Coonrod et al. 2001). In medaka, morpholino has been routinely used for gene knockdown experiments (M. Li et al. 2009; S. Porazinski et al. 2015).

1.4 Aims of the thesis

The first aim of my thesis study is to establish a new approach for easy and rapid identification of subtle allelic alterations produced by GE experiments. As an important technology in developmental genetic studies, GE has been developing rapidly recent years enabling efficient generation of subtle allelic alterations at site of interest using live embryos. Nevertheless, detecting and sequencing such alterations from potential individuals remain to be challenging since the efficiency of GE is not high enough for easy identification. In this thesis, I proposed and examined a PAGE

based HMA method which allows for the rapid and easy detection, quantification, and enrichment of subtle allelic alterations generated from GE experiments performed in medaka embryos.

Aside from manipulation of genes, the tracking of spatiotemporal dynamics of RNAs or specific cells is another important yet challenging task in developmental genetics. To date, *in situ* hybridization, which provides only static information, is the most frequently used method for visualizing RNA expression pattern or labeling specific cells. Only a few new methods achieved dynamic visualization of RNA with laborious procedure, and none of them is performed *in vivo*. Hence, the second aim of my thesis study is to develop a new approach for easy and specific visualization of RNA *in vivo*, which could be also used for specific cell tracking. In this thesis, I proposed and evaluated a molecular beacon probe using morpholino backbone which exhibited low toxicity, long term stability and high specificity for RNA visualization *in vivo*.

CHAPTER 2 Materials and Methods

2.1 General materials and molecular biology technique

2.1.1 Chemicals

In this study, chemicals were purchased from Sigma (MO, USA) unless otherwise indicated. Deionized water was collected from PURELAB Prima system from ELGA Lab (Singapore). Ultrapure water was collected from Milli-Q Water Purification Systems from Merck Millipore (Billerica, MA, USA).

2.1.2 Genomic DNA extraction

For DNA extraction from fin clips, whole fry or whole embryos, samples were collected from juvenile or adult fish and were incubated overnight at 55 ºC in 1.5-mL tubes containing 100 µL of lysis buffer (100 mM Tris-HCl pH 8.0, 200 mM NaCl, 0.2% SDS, 5 mM EDTA, and 100 µg/mL proteinase K, pH 8.0). Following heat-inactivation of proteinase K for 10 min at 70 ºC, DNA was precipitated with 2.5 volume of ethanol and was dissolved in 50 µL of TE (10 mM Tris-HCl, 1 mM EDTA, pH 8.0). DNA quantity was examined by WPA Biowave II UV/Visible Spectrophotometer from Biochrom (Cambridge, UK).

2.1.3 Polymerase chain reaction (PCR)

PCR was performed in Mastercycler from Eppendorf (Hamburg, Germany) with KAPA2G Fast PCR Kit from Kapa Biosystems (Wilmington, MA, USA) for indel detection, or with Phusion High-Fidelity DNA Polymerase from Life Technologies (Carlsbad, CA, USA) for plasmid construction. Three-step PCR protocol was followed. For KAPA2G polymerase, the protocol was: 2 min at 95 ºC, followed by 35 cycles of 15 s at 95 °C, 15s at annealing temperature of the primers calculated by

DNAman software from Lynnon Biosoft (San Ramon, CA, USA), and 30 s at 72 ºC, ending with a final extension of 5 min at 72 ºC. For Phusion polymerase, the protocol was: 30s at 98 ºC, followed by 33 cycles of 10 s at 98 °C, 15s at annealing temperature of the primers calculated by the Tm calculator of Thermo Fisher (https://www.thermofisher.com/), and 45 s at 72 ºC, ending with a final extension of 5 min at 72 ºC.

2.1.4 Agarose gel electrophoresis (AGE)

For agarose gel electrophoresis, 1.5% agarose gels were casted before each electrophoresis with 1×GelRed DNA stain from Biotium (Hayward, CA, USA). For each well, 2-5 μL sample was loaded together with Loading Buffer from Takara (Otsu, Japan). The gels were run at 5-8 V/cm in 1×TAE buffer (0.04 M Tris-base, 0.02 M acetic acid, 0.01 M EDTA, pH 8.0) with Horizontal Electrophoresis Systems from Bio-Rad (Hercules, CA, USA).

2.1.5 Polyacrylamide gel electrophoresis (PAGE)

For polyacrylamide gel electrophoresis (PAGE), 4-6 μL sample was loaded with loading buffer to 8% T, 3.3% C polyacrylamide gels. The electrophoresis was carried out in 1×TBE buffer (89 mM Tris base, 89 mM Boric acid, 2 mM EDTA) at 15V/cm for 1.5-2 h. After electrophoresis, gels were submerged in 1×TBE buffer containing 5×GelRed DNA stain for 15min to stain the DNA.

Both agarose and polyacrylamide gels were documented with GeneGenius Gel Doc System from Syngene (Cambridge, UK).

2.1.6 Purification of DNA from agarose gel or enzyme reaction solution

For purification of DNA from agarose gel, the DNA bands of interested were cut and retrieved under UV light, and then weighted. For purification of DNA from enzyme reaction solution, no prior steps was carried out. The purification was done using MN

NucleoSpin Gel and PCR Clean-up Kit from MACHEREY-NAGEL (Düren, Germany) according to its manufacturer's instruction.

2.1.7 Isolation of plasmid DNA

For DNA mini-preps, 3-5 mL of overnight bacterial culture in Lysogeny Broth (LB) medium (10 g tryptone, 5 g yeast extract, and 10 g NaCl in 1 L deionized water, pH 7.0) were used. For sequence screening, an alkaline lysis protocol was followed to extract plasmid DNA as described (Birnboim and Doly 1979). Alternatively, for construction of plasmids, QIAprep Spin Miniprep Kit from Qiagen or MN NucleoSpin Plasmid Kit from MACHEREY-NAGEL were used according to their user's guide.

For DNA midi-preps, MN Nucleobond Xtra Midi from MACHEREY-NAGEL kit was used.

2.1.8 Isolation of total RNA and first strand cDNA synthesis

To extract total RNA, more than 30 mg of fish tissues, or more than 20 fish embryos, which showed high expression of the target gene, were used for each sample. RNA was extracted by TRIzol reagent from Life Technologies following its user's guide. Briefly, tissues or embryos were first homogenized in 1mL TRIzol reagent on ice; then 0.2 mL of chloroform was added to the homogenized sample, and the sample tubes were capped and shaken vigorously; after centrifugation, take the aqueous phase out of the tube for isopropanol precipitation of the RNA.

First strand cDNA synthesis was performed using RevertAid First Strand cDNA Synthesis Kit from Life Technologies following the user's guide.

2.1.9 Cloning and sequencing

TA-cloning method was used to clone DNA segments from PCR products (Sambrook and Russell 2001). For PCR products from Phusion polymerase, 3' A overhangs were

added prior to TA-cloning by purifying the PCR product and incubating with KAPA2G Fast PCR polymerase along with its buffer for 20 min at 72 °C. TA-cloning was performed with pGEM-T Easy vectors from Promega, and the ligation products were transformed to commercial competent cells HIT-DH5alpha High 10^8 from BioAspect (Toronto, ON, Canada) according to its user's guide, or to self-prepared competent cells (from strain TOP10F' from Invitrogen) according to Inoue's protocol (Sambrook and Russell 2001). Single colonies were pick for inoculation in 3-5 mL LB medium followed by mini-prep of plasmid DNA. The extracted plasmids were subjected to cycle sequencing of the interested cassettes. BigDye Terminator Cycle Sequencing Kit from Life Technologies was used for the cycle sequencing as instructed in the user's manual. The cycle sequencing products were purified with PureSEQ-MP magnetic beads from Aline Biosciences (Woburn, MA, USA) before sequence analysis on the 3130xl Sequencer (Applied Biosystems). Sequence alignment were done by Vector NTI from Life Technologies and DNAman software from Lynnon Biosoft (San Ramon, CA, USA).

2.2 Fish maintenance and microinjection of fish embryos

2.2.1 Fish strains and maintenance

Work with fish followed the guidelines on the Care and Use of Animals for Scientific Purposes of the National Advisory Committee for Laboratory Animal Research in Singapore and was approved by this committee (permit number 27/09). Briefly, for embryo collection, medaka adults (strains *af*, *HdrR* and *HdrR II*) were maintained in ~20 L tanks of static water (fresh dechlorinated tap water with hardness under 100mg/L $CaCO_3$ equivalent) under an artificial photoperiod of 14-h/10-h light/darkness at 26 - 28 °C. 15-20 females with 4-5 males were kept in each tank. Brine shrimp was fed three times daily. Freshwater snails were kept in each tank for

better water quality. One-third of water in each tank was changed each week during the removal of debris from the bottom of the tank. Water pH was monitored weekly and was maintained to be 6.5 - 7.0. For detail information about medaka maintenance, a monograph (Kinoshita et al. 2009) can be referred to.

2.2.2 Microinjection

Microinjection of medaka embryos was performed as described (S. R. Porazinski, Wang, and Furutani-Seiki 2010). Briefly, to prepare the medaka embryos for microinjection, male and female fish were separated one night before the microinjection and were mixed to allow courtship and embryos spawning right before the experiment. The embryos were collected soon after spawned. They were then separated and arranged in agarose grooves (made with 1.5% agarose using customized moulds of 0.8 × 1 mm ridges) submerging in 1x Yamamoto Ringer's solution (NaCl 7.5 g, KCl 0.2 g, $CaCl_2 \cdot 2H_2O$ 0.2 g, and $NaHCO_3$ 0.02 g in 1000 mL H_2O, pH 7.3). These embryos were rotated with their cytoplasm facing upwards. Glass capillary needles were pulled with a Micropipette Puller P-87 (Sutter Instrument, USA). Solution to be injected into the embryos was loaded to the glass capillary needles by first heating the needles to expand the air inside and then cooling it so that air contraction could bring the solution in. After loading the solution, the needles were assembled to the Mechanical Manipulators by Leica Microsystems (Wetzlar, Germany) with a self-made micro-injector using mineral oil as pressure transmitter. Then, the tips of the needles were opened through touching slightly on hard surfaces and microinjection could begin. Usually, 0.4 - 1 nL of the solution would be injected into the cytoplasm of 1-cell or 2-cell stage embryos. After injection, the embryos were incubated at 28 ºC with embryo rearing medium (NaCl 1.00 g, KCl

0.03 g, $CaCl_2 \cdot 2H_2O$ 0.04 g, $MgSO_4 \cdot 7H_2O$ 0.16 g, and optionally, Methylene blue 0.0001 g in 1000 mL H_2O, pH7.3).

2.2.3 Photography of embryos

Embryos were immobilized on agarose grooves or with 3% methylcellulose during photography. Pictures were captured with QImage CCD and QCapture Pro software from QImaging (Surrey, BC, Canada) under Fluorescence Stereo Microscope M205FA from Leica.

2.3 Detection of subtle DNA alterations by PAGE-HMA

2.3.1 Gene editing by Zinc finger nucleases (ZFNs) & TALENs

Embryo samples of gene editing by ZFN were provided by Dr. Guan Guijun. Expression vectors pZN1gsdf and pZN2gsdf (Chen et al. 2012) were supplied commercially (Toolgene, South Korea). They encode a pair of ZFNs that target the first exon of the medaka *gsdf* gene. After linearization with ApaI, both vectors were used for mRNA synthesis by using the T7 RNA polymerase (Invitrogen) as described (M. Li et al. 2009). The pair mRNAs were mixed at 1:1 ratio and microinjected at 20~100 ng/μL into embryos of *af* and *HdrR* medaka at the 1-cell stage as described (M. Li et al. 2009). The injected embryos were analyzed 3 dpf or allowed to develop into adulthood, where adult fin clips were extracted for gene editing event screening.

Embryo samples of gene editing by TALENs was provided by Dr. Wang Tiansu. Expression vectors pTN1dnd and pTN2dnd was designed and constructed as described (T. Wang and Hong 2014). Synthetic mRNAs transcribed from these vectors were microinjected into 1-cell stage medaka *HdrR* embryos. The injected embryos were raised to adults and genomic DNA samples from their adult fin clips were extracted for gene editing event screening.

2.3.2 Genome editing by CRISPR-Cas

Plasmids of CRISPR-Cas gRNA (ID:46759) and Cas9 nuclease (ID:47929) from Chen and Wente Lab (Jao, Wente, and Chen 2013) were obtained from Addgene plasmid repository (Cambridge, MA, USA). The medaka *nanog* mRNA sequence was obtained from NCBI (GenBank ID: NC_019878.1). Two targeting sites of the *nanog* first exon were selected manually with GG at both ends. Whole medaka genome BLAST of the gRNA sequences were performed to evaluate off-target effect. No site can be found with >15 bp homolog followed by NGG. The CRISPR-Cas system preparation protocol used can be found at http://www.addgene.org/crispr/Chen/. Briefly, two mostly complementary oligos, which contain the targeting site and the overhangs for ligation, were synthesized and annealed to form double stranded segments. In my case, oligo sets gRNA OlNanog 183 and gRNA OlNanog 319 were used (see Appendix Table 1 or Fig. 3-12 A). One-step digestion and ligation of the double stranded segments and the gRNA vector was carried out using T4 ligase and three endonucleases: BsmBI, BglII and SalI. This reaction was performed in the Mastercycler with the following program: two-step cycles of 20 min at 37 °C and 15 min at 16 °C for three cycles, followed by 10 min at 37 °C, 15 min at 55 °C, and 15 min at 80 °C. The ligation products were transformed to E. coli and colonies were picked and the inserts were sequenced for the desired gRNA vectors. These correct gRNA vectors and the Cas9 nuclease vector were linearized, and their RNA was transcribed from the linearized products using T7 or SP6 mMESSAGE mMACHINE Transcription Kit from Life Technologies. For microinjection of the Cas9 mRNA and gRNA, 4 - 8 nl of 100 ng/µL gRNA with 100 ng/µL Cas9 mRNA was co-injected into 1-cell stage medaka *HdrRII* embryos. The injected embryos were analyzed or allowed to develop into adulthood.

2.3.3 Plasmid construction for the proof of concept studies

WT1 and D18 are 321 bp and 303 bp in length, which were PCR-amplified by using primers nanog WT1 (Table 1) from two plasmids containing the medaka *nanog* cDNAs. pWT2 contains a 346-bp insert of the first exon of medaka *gsdf* gene. The insert was PCR-amplified by using primers gsdf WT2 (Table 1) and cloned in pGEM-T Easy vector. pD1, pD6, pD7, pA18 and pA3D7 were similar to pWT2 except for deletions (D) or additions (A) of 1, 6, 7 and 18 bp in the insert. These deletions or additions were generated from ZFN-mediated gene disruption in medaka embryos (see 3.4.1).

2.3.4 Heteroduplex formation

When DNA mixtures containing two or more different sequences is used as templates for PCR, heteroduplexes are formed automatically during cycles of denaturing and re-annealing (Ruano and Kidd 1992). However, in my thesis, to assure full heteroduplex formation or to generate heteroduplexes between PCR products from two separate reactions, a denaturing and re-annealing procedure was performed. PCR products or mixture of PCR products from separate reactions were heated to 94 °C for 3 min and slowly returned to room temperature to form heteroduplexes (Highsmith et al. 1999).

2.3.5 T7 Endonuclease I AGE assay

For the traditional TAGE detection method, 3 μL of PCR product was subjected to digestion with 0.3 U of T7 endonuclease I (New England BioLabs) in a 5 μL volume at 37 °C for 1.5 h. Then, 5 μL digested products and3 μL undigested PCR product as control were mixed with 5× loading dye, respectively, and were loaded on the 2% (w/v) agarose gel with 0.05 μg/mL Ethidium Bromide to run the electrophoresis under 130 V for 15 min.

Table 1. Sequences of the primers used.

Name	Forward (5'-3')	Reverse (5'-3')
Nanog WT1	ATGGTTGAGTCCCAATCTTTTG	TCAATATCGCTCTGAAACCCAG
Nanog 445	ATGGCGGAGTGGAAAACTCAGGTC	GGCAGAACTTGAAGAAGAACTCGC
Nanog 305	TCGCCACTCCTTTCTCCTCC	GAAAAGGATCGTTGCAATTCGC
Nanog 372	GGTTCCCAGGCCCCGGGCTAGCAG	CTATGACCAGCATTTGCAGAGTTAGC
Nanog 572	GCATTTGTTGCCAGTCTCATACATACATG	GGCAGAGTGACTCATTCCACTGATG
Gsdf WT2	GGTTCCCAGGCCCCGGGCTAGCAG	GAAAAGGATCGTTGCAATTCGC
Dnd WT	TTTACATTTCGTGTTCAGTGTGG	GTGCCGTATTTGGCGTAAG
gRNA OlNaong 319	TAGGCGAAAACTGCACGAGA	AAACTCTCGTGCAGTTTTCG
gRNA OlNanog 183	TAGGAGTCCCCACCACGCAGCC	AAACGGCTGCGTGGTGGGGACT

2.3.6 PAGE-HMA

For PAGE-HMA, 4 μL PCR products were separated on 8% polyacrylamide gels in 1×TBE buffer using a Mini-Protean electrophoresis unit (Bio-Rad Laboratories) at 100 V for 2 h in room temperature. Gels were submerged in 1×TBE buffer containing 0.05 μg/mL ethidium bromide for 20 min. Gels were documented on a bioimaging system (Vilber Lourmat).

2.3.7 Densitometry analysis of PAGE gels

The documented gel pictures were analyzed by the Gel-Pro Analyzer software from Media Cybernetics (Rockville, MD, USA). During the analysis, frames for separating bands were drawn manually according to known band position from Fig. 3-3. Percentages of all bands from one lane were given by the software after framing. Band intensity of each allele was calculated by summing up the percentages of bands that containing the allele.

2.3.8 Gel recovery and PCR

For gel recovery, a protocol was used based on the crush and soak method (Sambrook and Russell 2006). DNA bands on the PAGE gel were cut under UV light, and were then crushed and soaked in 10~20 μL of TE. After incubation overnight at room temperature, 1 μL of the supernatant containing DNA was used for PCR (see above).

2.3.9 Identification of indels from gene edited individuals

To identify indels from individual fish or embryos, DNA from the candidate fish or embryos were extracted for PCR amplification of segments flanking the gene editing target sites. The PCR products were analyzed with PAGE-HMA and heteroduplex bands of interest were recovered for sequencing as described above (section 2.1.9). For ZFN-mediated gene editing of *gsdf*, primers gsdf WT2 were used. For TALEN-

mediated gene editing of *dnd*, primers dnd WT were used. For CRISPR-Cas-mediated gene editing of *nanog*, primers nanog 455 were used.

2.4 Visualization of RNA *in vivo*

2.4.1 Design and preparation of probes

2.4.1.1 Morpholino molecular beacons design

Three morpholino molecular beacons (MOMBs) with stem length of five (MOMB-5), seven (MOMB-7), and ten (MOMB-10) were designed manunally along with their target sequence (see Appendix Table 2). Shared-stem design was adopted in which one arm of the stem complements the target site. Secondary structures of these MOMBs were checked by Quikfold application from the DINAMelt Web Server http://mfold.rna.albany.edu/ (Markham and Zuker 2005; Markham and Zuker 2008) with temperature set at 28 °C and other settings as default for DNA folding. Both probes show regular stem-loop structure as desired. A DNA molecular beacon (DNAMB) was also designed with the same sequence as MOMB-7. MOMB and DNAMB were purchased from Gene Tools (Philomath, OR, USA) and Integrated DNA Technologies (Coralville, IA, USA), respectively.

2.4.1.2 Preparation and purification of probes

All probes ordered were delivered as dry powder and were dissolved as 1 mM stock solution in ultrapure water upon arrival. To prepare the MOOM displacing probes, 10 μM reporter strand and 12 μM quencher strand were mixed in phosphate buffer (PB: 1 mM Na_2HPO_4 + KH_2PO_4 pH 7, 100 mM NaCl, 3 mM $MgCl_2$) for overnight. Before use, MOMB probes were purified twice with ultrapure water through Amicon Ultra-15 Centrifugal Filter 3kDa NMWL by Merck Millipore to wash out leftovers of dissociative fluorophores, which result from the imperfect purification of MO products from Gene Tools.

2.4.2 Characterization of MBs *in vitro*

2.4.2.1 Fluorescence emission spectra and hybridization kinetics

Fluorescence emission spectra and hybridization kinetics analysis were performed using CYTATION 3 Cell Imaging Multi-Mode Reader from BioTek (Winooski, VT, USA) with pure grade black microplates from BRAND (Wertheim, Germany). Fluorescence emission spectra of 500 nM DNA-MB or MO-MB in 200 μL solution was acquired from 510 nm to 650 nm with 1 nm step with monometer. Hybridization kinetics were recorded with GFP filter (485/528) with sampling intervals of 15 s.

2.4.2.2 Thermal denaturation analysis

Thermal denaturation analysis were performed with final concentration of 500nM DNA-MB or MO-MB in 20 μL solution using StepOnePlus Real-Time PCR System from Applied Biosystems (Waltham, MA, USA) or CFX96 Touch™ Real-Time PCR Detection System from Bio-Rad. The solution was heated from 26 °C to 90 °C, and the fluorescence signal was collected for every 2 °C increase after 10 s holding time.

2.4.2.3 Endonucleases and DNA binding protein assays

RNase H and T4 Gene 32 Protein for DNA binding were purchased from New England BioLabs (Ipswich, MA, USA). RQ1 DNase were purchased from Promega (Madison, WI, USA). All assays were performed with CYTATION 3 Cell Imaging Multi-Mode Reader at of each at 37 °C. Fluorescence intensity was recorded every 15s during the assays with the GFP filter (485/528).

Rnase H digestion assay

Perfect match RNA oligos were added with final concentration of 2.5 μM to 200μL reaction system containing 500 nM DNA-MB or MO-MB and 1× reaction buffer as provided. After hybridization for 30min, fluorescence intensity of each solution was recorded for 5 min. Then, 15 U Rnase H was added to each solution, and the

fluorescence intensity changes were monitored for 2 h.

DNase digestion assay

Fluorescence intensity of 500nM MB or MO-MB in 40mM phosphate buffer with 10mM $MgCl_2$ and 1mM $CaCl_2$ was recorded for 5 min, followed by the addition of 2 Units of RQ1 DNase and the record of fluorescence intensity changes for 2 h.

Single-stranded DNA binding protein (SSB) binding assay

Fluorescence intensity of 500nM MB or MO-MB in 200 μL 1×NEB buffer 4 (provided by manufacturer, 50mM Potassium acetate, 20mM Tris-acetate, 10mM Magnesium acetate, 1mM DTT, pH 7.9) was recorded for 5 min, followed by the addition of 5-fold excessive SSB T4 Gene 32 Protein (100 μg) and the record of fluorescence intensity changes for 2 h.

2.4.3 Characterization of MBs *in vivo*

2.4.3.1 Toxicity assays *in vivo*

Based on the OECD Fish Embryo Acute Toxicity Test (OECD 2013) which uses zebrafish embryos, *in vivo* toxicity experiments of the probes on medaka fish were designed and performed as followed. 50 μM or 250 μM probes were mixed with 5 mg/mL Texas Red and were injected to 1-cell stage *HdrRII* embryos. Texas Red of 5 mg/mL was also injected as control group to evaluate microinjection caused damage. Right after injection, the injected embryos were screen under fluorescent microscope to remove unsuccessful injection samples. To evaluate the toxicity of probes, abnormality and death rate of each injection groups on 1 dpf, 5 dpf and 3 dph were recorded. In this evaluation, embryos with abnormal epiboly, somite, head and cardiovascular system development would be recorded as abnormal, while embryos without body axis on 1dpf and without heartbeats from 5dpf onwards would be recorded as dead.

2.4.3.2 Stability assays *in vivo*

2.4.3.2.1 Stability assay via fluorescence intensity

To compare the stability of DNA probes and morpholino probes *in vivo*, 50 μM DNA probes, 100 μM MOMB or MOOM probes were mixed with 5 mg/mL Texas Red respectively and were microinjected to 1-cell stage *HdrRII* embryos. Pictures of the green and red channels of the injected embryos were taken every 15 min to monitor the degradation rate of each probe. All pictures were taken under the same exposure condition: 36 bit, 700ms exposure time for green channel and 200ms for red channel. The original image files were cropped into 100×100 pixels, remaining the center parts of the cell with uniform fluorescence intensity. The cropped images were subjected to brightness analysis using a Python script to get a green/red ratio by averaging brightness value of all pixels (see Appendix Python scirpt 1). Background noise was subtracted using a blank background image as reference.

2.4.3.2.2 Stability assay via LC-MS

To further assess the stability of MO-MB *in vivo*, LC-MS was used to determine the existence of MO-MB. To this end, 250 μM MO-MB was injected to 1-cell stage *HdrRII* embryos. Around 40 embryos were dechorionated by forceps at 6 h and 24 h after injection with the cell mass separated from the yolk and subjected to overnight lysis by 0.1 mg/mL Proteinase K in 0.05% SDS. The cell lysis was centrifuged at 14000 g for 10 min and 50 μL supernatant was taken for LC-MS analysis. Pure MO-MB of 50 μM was used as reference and uninjected 6 hpf embryos were used as negative control.

The HPLC System used was Dionex UltiMate 3000 and the MS system Bruker MicrOTOF-QII. Phenomenex Aeris Widepore 3.6u XB-C8 was used as the HPLC column. Two buffers were use as followed: Buffer A: H_2O + 0.1% formic acid;

Buffer B: ACN + 0.1% formic acid. Flow rate was set at 0.2 mL/min. and the effective gradient program was set as followed: 0-5 min 15% B, 6-10 min 60% B, 11-25 min 100% B.

2.4.3.3 Specificity assays *in vivo*

For *in vivo* specificity assays of MOMB probe, 200 μM probes were injected to 1-cell stage *HdrRII* embryos. Injected embryos were checked under fluorescence microscope to confirm successful injection. The embryos were then incubated at 28°C for three hours to reach 16-cell stage. One single cell of the 16-cell embryos was injected with 10 μM perfect match target or 2-mismatch target mixed with 5 mg/mL Texas Red as indicator. The embryos were pictured 15min after the injection of the targets.

Table 2. Sequences of the molecular beacons and targets.

Name	Backbone	Sequence (5'-3')
MOMB-5	Morpholino	DABCYL-GCCAGGATAGAATAGTCACACTGGC-FAM
MOMB-7	Morpholino	DABCYL-GCGCGCTATACACGACACAGCGCGC-FAM
MOMB-10	Morpholino	DABCYL-GGAGCCAGCGCCTTTCGCTGGCTCC-FAM
DNAMB	DNA	DABCYL-CGCGCGACACTCCAGAAATCGCGCG-FAM
MOMB-5 target	RNA	UGUGACUAUUCUAUCCUGGC
MOMB-7/DNAMB target	RNA	UGUGUCGUGUAUAGCGCGC
MOMB-7/DNAMB target 2MM	RNA	UGUGUGGUGUAUAGGGCGC
MOMB-10 target	RNA	GCGAAAGGCGCUGGCUCC

CHAPTER 3 Results and discussion I:

Detection of subtle DNA alterations by PAGE-HMA

3.1 Rationale and experimental design

As reviewed in the introduction, for embryo microinjection based gene editing (GE) experiments, one major challenge is to identify the individuals as well as the sequences of successful gene editing events with minor indels from hundreds of candidates. We were in need of a convenient and labor-saving method to solve this problem. Therefore I developed a polyacrylamide gel electrophoresis (PAGE) based heteroduplex motility assay (HMA) detection procedure termed PAGE-HMA, which draws upon the formation of heteroduplexes during PCR of the GE target site when indels exist.

After gene editing process, genomic DNA containing indels will generate three types of double-stranded PCR products, namely, wild-type (WT) homoduplex (Hm), mutant (MT) Hm, and WT-MT heteroduplex (Ht). These products are naturally generated due to the repeated cycles of denaturing and annealing, and do not show a visible difference on agarose gel electrophoresis (AGE). A traditional assay T7EI-AGE (TAGE) utilizing T7 endonuclease I, which can digest Ht DNA fragments at the mismatch site, is routinely used to identify the Ht. The TAGE assay can serve as a control experiment to verify my new method.

In the PAGE-HMA, Ht and Hm can be directly separated on native PAGE due to the differences in conformations. According to the principle of HMA, Ht shift slower than Hm on PAGE; therefore, Ht bands can be found above Hm bands, as an unambiguous

indicator of the existence of indel. In addition, Ht bands could also be separated from Hm bands due to their differences in molecular weight. Moreover, all PCR products remain intact after PAGE, unlike TAGE that cut the target PCR product. Hence, once the Ht band is visualized on the poly-acrylamide gel, the DNA fragments contained within could be recovered by excising the band. This feature allows us to enrich the Ht through a second round of PCR prior to cloning and sequencing, which highly increases the percentage of indel-containing colonies, lessen the workload required for detecting an indel event.

As depicted in Fig. 3-1, I proposed a workflow for identifying gene editing events based on the PAGE-HMA method. In the work flow, medaka embryos at 1-cell stage will be injected with engineered nucleases, for example, zinc finger nucleases (ZFN), transcription activator-like effector nucleases (TALEN) and CRISPR-Cas (CRISPR: clustered regularly interspaced short palindromic repeats; Cas: CRISPR-associated). The injected embryos will be raised to adult. Genome DNA of each adult will be extracted individually from the adults through tail fin clip. Samples containing indels at the target sequence would form Ht with WT alleles during PCR. These Ht can be revealed by PAGE and be further recovered by cutting, crushing and soaking the gel containing Ht as described in section 2.3.8. The recovered Ht will be subjected to another round of PCR for amplification followed by cloning and sequencing to confirm the indel profile. Adult fish that containing frame shifted indel can be selected as founder for breeding. Their progenies would experience the same screening process to establish or maintain the knock-out fish strain.

Before this proposed method was applied to real GE experiment, proof-of-principle experiments were carried out first to evaluate three parameters: sensitivity, discrimination of multiple alleles, and efficiency of enrichment. Accordingly, a

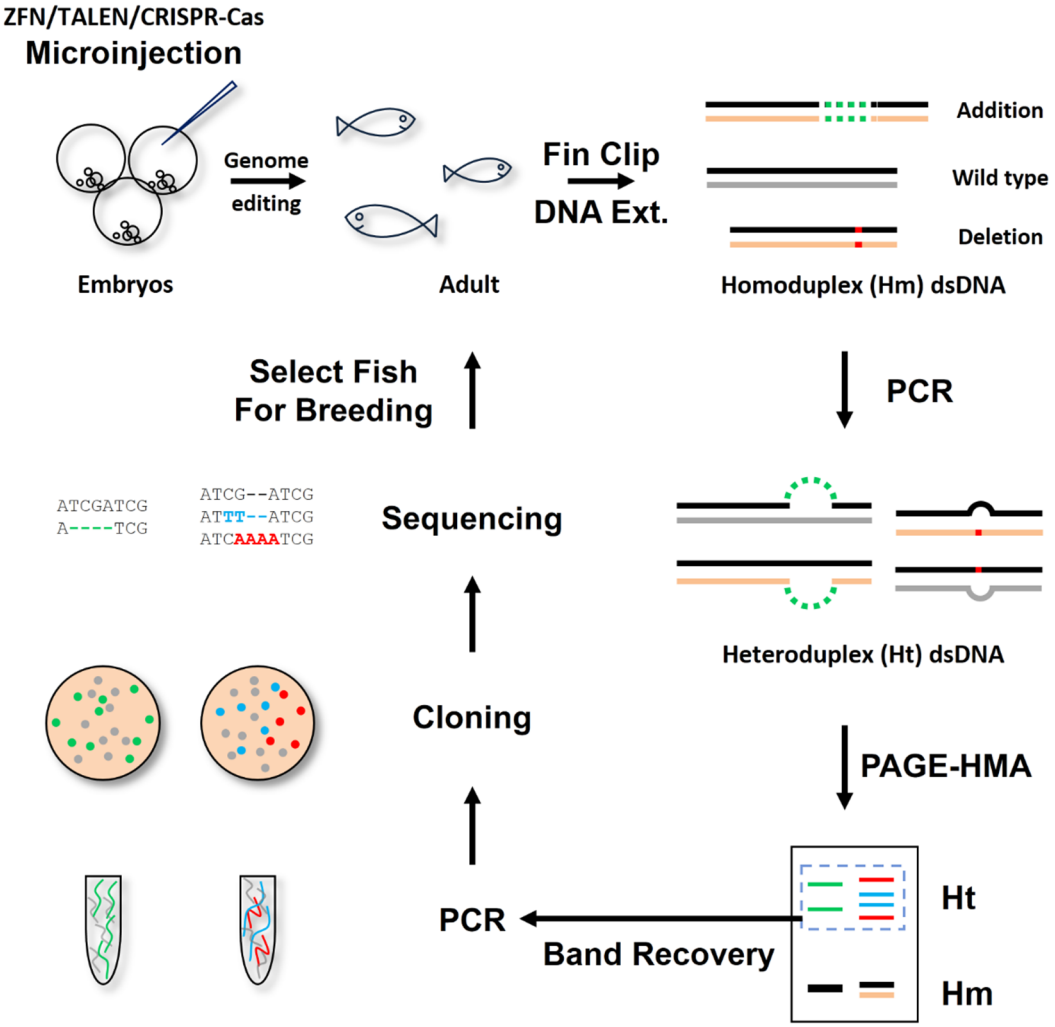


Figure 3-1. Workflow of identifying genome editing events (In collaboration with X. Zhang, see DNA Res. 2012 Oct;19(5):423-33)**.** In our routine genome editing experiments, embryos injected with engineered nucleases at 1-cell stage are raised to adulthood. Genome DNA is extracted by fin clip and is subjected to PCR and PAGE-HMA. Ht bands on PAGE indicating successful genome editing events. These bands can be further recovered, followed by cloning and sequencing for confirmation of the event. The individuals with positive genome editing events would be selected as founder for breeding. Their progenies would go through the same screening process. Hm, homoduplex; Ht, heteroduplex; PAGE-HMA, polyacrylamide gel electrophoresis based heteroduplex mobility assay.

biallelic model containing two slightly different alleles were used to test sensitivity of the method; a multi-allelic model containing five types of alleles were used to test its ability to discriminate different alleles; and finally, cloning and sequencing of the recovered alleles were performed to assess its enrichment efficiency.

3.2 Proof of principle experiments

3.2.1 Biallelic model

Procedures and the usefulness of TAGE and PAGE-HMA were first established and evaluated for Ht detection in a biallelic model system. To this end, two plasmids, WT1 and D18, which contain an insert of 321 and 303 bp as the WT allele and 18-bp deletion allele (thus D18) of the medaka *nanog* gene, were linearized and mixed to form a series of dilution. To mimic the low percentage of indel-containing alleles in real gene editing experiments, D18 was diluted by WT1 at a factor of 2 to 256. This serial dilution mixtures were used as the templates for PCR (100 pg templates in 25 μL PCR reaction), and the PCR products were subjected to both TAGE and PAGE-HMA detection.

The result showed that with TAGE, Hm WT1 and D18 amplicons remained intact and appeared as bands that were slightly different in size (Fig. 3-2 A). When mixed at a 1:1 ratio, WT1 and D18 without T7 cleavage produced a double band. Upon T7 cleavage, two distinct bands with a similar intensity could be seen. When mixed at serial ratios, D18 remained detectable at the ratio of 1:16, indicating a sensitivity level of ~6.3% for TAGE in Ht detection. Since the lower limit of each detection method was the focus, results with dilution factor 4 and 8 are not shown. T7 digestion usually leads to reduced band intensities when detecting Ht, because T7 cuts the Ht amplicons into two smaller fragments which might not be the same size and thus might not form a condensed single band. Therefore, it is very difficult for the TAGE to detect indels

of low percentages, which could produce false negative results during gene editing screening.

As for the PAGE-HMA detection, when mixed at a 1:1 ratio, WT1 and D18 mixture showed two Hm bands and two Ht bands (Fig. 3-2 B). Two Hm bands lay at the same position as WT1 and D18 homoduplex. The distance between the two Hm bands came from their size difference. Two Ht bands lay above the Hm bands due to their variation of configuration. As illustrated in Fig. 1, each MT allele could form two different configurations of Ht fragments with the WT allele depending on the involvement of either upper- or lower- strands (Nataraj et al. 1999). These Ht fragments migrated much slower on PAGE than Hm fragments, forming bands with a clear distance from Hm DNAs. Using Ht bands as the indicator, D18 can be unambiguously detected even at a dilution factor of 256, which lead to a sensitivity level of ~0.4% for PAGE-HMA in Ht detection, 16-fold more efficient than TAGE. This higher sensitivity of PAGE-HMA stemmed from both the large distance between Ht and Hm bands and the non-smear band profile of Ht on PAGE. Moreover, compared with only one band as indicator on TAGE, PAGE-HMA showed two Ht bands, which also contributes to their easy identification.

Taken together, using the biallelic model, I showed that the PAGE-HMA method can detect MT alleles at sensitivity of 0.4% in the absence or presence of genomic DNA, which was 16-fold more sensitive compared with the widely used TAGE method. Moreover, Ht fragments on TAGE formed a faint band embedded in a smear at a dilution factor of 16, while on PAGE-HMA, Ht exhibited as sharp bands with little background even at a dilution factor of 256.

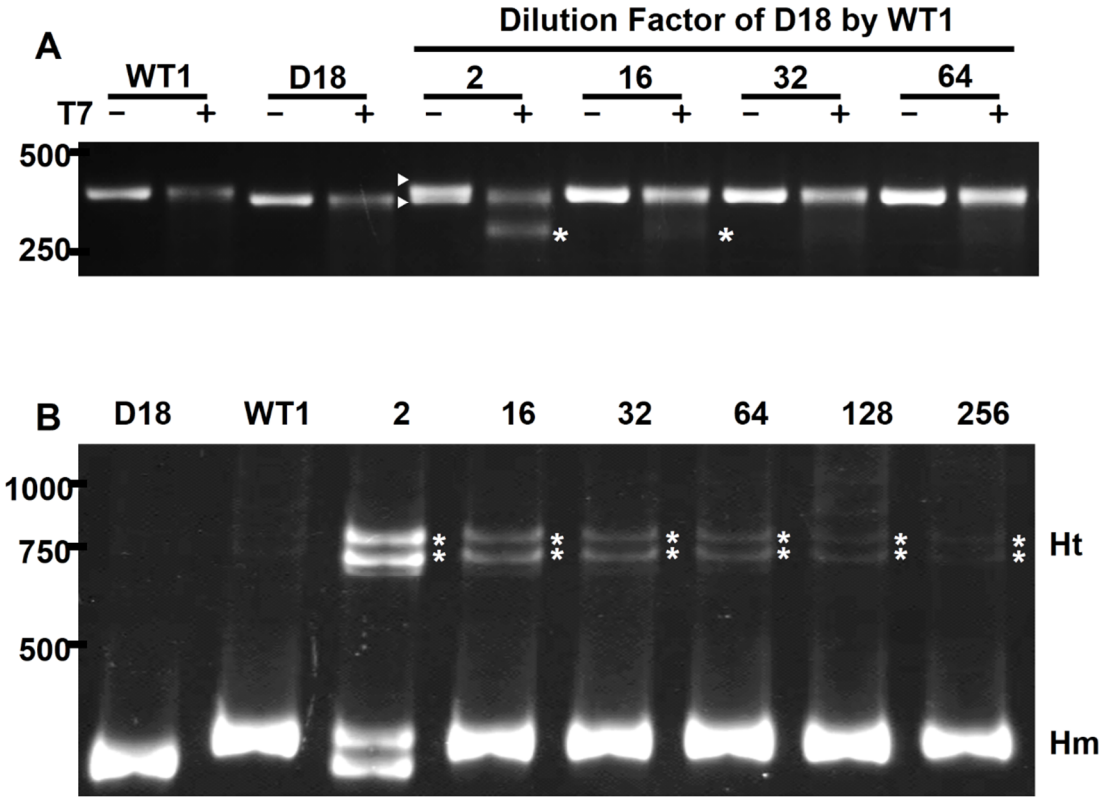


Figure 3-2. Heteroduplex detection by TAGE and PAGE-HMA (In collaboration with X. Zhang, see DNA Res. 2012 Oct;19(5):423-33 & X. Zhang's Thesis)**.** Plasmids containing a wild-type allele (WT1) and a mutant allele with a 18-bp deletion (D18) were mixed to reach the indicated dilution factor for PCR. (A) T7 endonuclease I digestion followed by agarose gel electrophoresis (TAGE) profile of the PCR product. D18 with a dilution factor of 16 was detected. (B) Direct PAGE-HMA of the PCR product. D18 with a dilution factor of 256 was detected as heteroduplex bands. Size markers in bp are indicated to the left. Arrowheads denote a double band. Asterisks highlight heteroduplexes. Hm, homoduplex; Ht, heteroduplex; PAGE-HMA, polyacrylamide gel electrophoresis based heteroduplex mobility assay; T7, T7 endonuclease I; TAGE, T7-digestion followed by agarose gel electrophoresis.

3.2.2 Multi-allelic model

After evaluating the PAGE-HMA method with the biallelic model, I further conducted experiments with a multi-allelic model system to test the ability of PAGE-HMA in detecting multiple alleles. In the practice of gene editing experiments, additions and deletions may occur independently in different cells and at different stages of development, which leads to the production of various different MT alleles. To stimulate such a scenario, a set of five different alleles were produced within exon 2 of the medaka *gsdf* gene via ZFN-mediated GE in medaka embryos (see 3.6.1). They were WT2, A18, D7, D6 and D1, where A represents addition (insertion) of nucleotides and D represents deletion (Fig. 3-3 A). WT2 of 320 bp in length is the WT allele of *gsdf* (WT1 described above is the WT allele of the medaka *nanog*). These allelic sequences were cloned to pGEM-T Easy vector, and linearized plasmids were mixed at various ratios as described in Fig. 3-3 B. As expected, PCR produced a single Hm band in PAGE-HMA when a single plasmid was used as template (Lanes 1~5, Fig. 3-3 B). When two plasmids were used, Ht bands above the Hm bands could be found (Lanes 6, 7, 9, 11, Fig. 3-3 B), even when the sequence difference is as tiny as 1 bp deletion (Lane 6 asterisk, Fig. 3-3 B). Most importantly, when multiple plasmids were used, they could be resolved as various distinct Ht bands (Lanes 8, 10, 12, 13 Fig. 3-3 B), and the same Ht seated on the same position on different lanes, as linked by dash lines in Fig. 3-3 B. The number of alleles in a mixture did not affect their Ht profile. For example, on Lane 12 Fig. 3-3 B, the mixture consisting of as many as four different alleles still replicated the same pattern as seen on Lane 7-11 Fig. 3-3 B. Moreover, the number of alleles in a mixture did not affect the sensitivity of PAGE-HMA detection, as shown on Lane 13 Fig. 3-3 B where each MT allele was

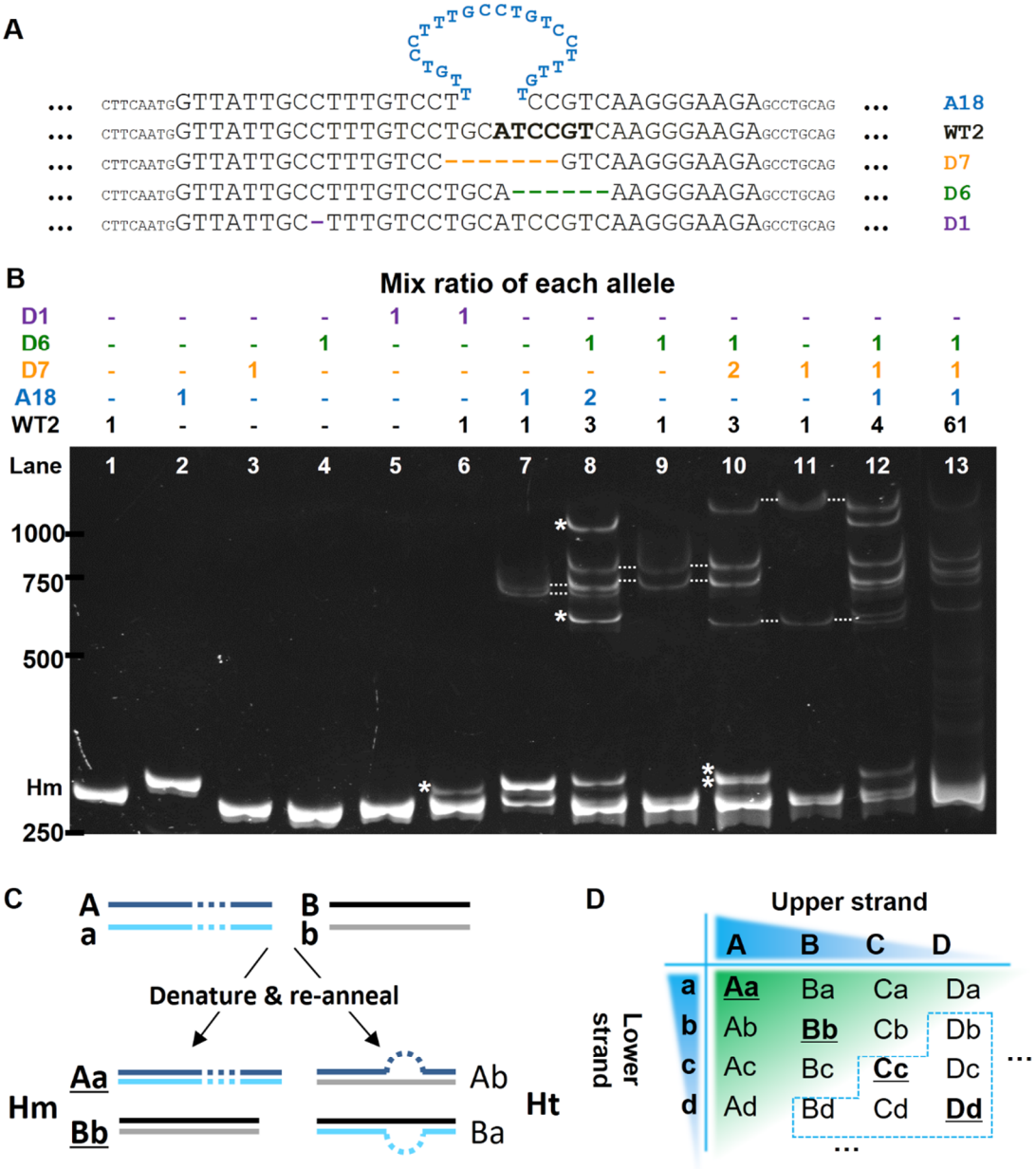


Figure 3-3. PAGE-HMA detection of multiple alleles (In collaboration with X. Zhang, see DNA Res. 2012 Oct;19(5):423-33 & X. Zhang's Thesis)**.** (A) Partial sequences of 5 different alleles of the medaka *gsdf* gene used in the multi-allelic model. Bold letters indicate the cleavage site of ZFN. (B) PAGE-HMA profiles of PCR products from different mixtures of the alleles. Mix ratios of the alleles are indicated above each lane. Different mixtures exhibit different profile with bands from the same alleles appear on the same position. Dash lines link these bands. Asterisks highlight heteroduplexes. (C, D) Correlation between the number of alleles and total number of different PCR products which become different bands on PAGE. (C) The formation of homoduplexes and heteroduplexes during PCR. (D) All possible combinations of heteroduplexes for four different alleles. The intensity of the blue background depicts the positive correlation between the concentrations of each alleles and their corresponding heteroduplexes. Bold and underlined letters denote homoduplexes. Capital letters, upper strands; lower case letter, lower strands; Hm, homoduplex; Ht, heteroduplex.

diluted for 64 times and was still detectable as Ht bands with the same profile as on Lane 12 Fig. 5B.

To confirm that these Ht bands were not from artificial non-specific PCR products, PCR products manifesting single band in Lane 1-5 of Fig. 3-3 B were mixed at the same ratio as indicated in Fig. 3-3 B and then were subjected to a heteroduplex formation procedure. PAGE-HMA profiles (Fig. 3-4) of these mixtures show exactly the same patterns as in Fig. 3-3 B.

The ability of PAGE-HMA to separate different Ht lies on the principle of HMA. Ht with different conformation could have different electrophoresis speed (Nataraj et al. 1999). Therefore, in my case, indels could result in various Ht as shown in Fig. 3-1, forming loops of different lengths, which were naturally separated on native PAGE. Moreover, besides the loop lengths, sequence differences of each Ht could also conduce to their separation on PAGE.

To better understand the PAGE-HMA profile of a mixture of alleles, I established the general correlation between the total number of Ht bands on PAGE-HMA and the total number of alleles in a mixture. The correlation is not a simple linear positive one. As revealed by Fig 3-3 B, two alleles result in two Ht bands and in two Hm bands which might not be distinguishable (Lane 7, 9, 11), yet three alleles produce six Ht bands and three Hm bands (Lane 8, 10). Note that the additional two new Ht bands were not seen in the subset of their mixture (Asterisks). This phenomenon can be seen also in Fig. 3-4 B. To explain this, I proposed that, as illustrated in Fig. 3-3 C, every two alleles in the mixture could form two Ht. Therefore, in a mixture containing four alleles, as depicted in Fig. 3-3 D, there would be 16 types of configurations, among which are 12 Ht and 4 Hm. In general, the number of Ht would be the number of all possible combinations between every pairs of alleles minus the number of Hm.

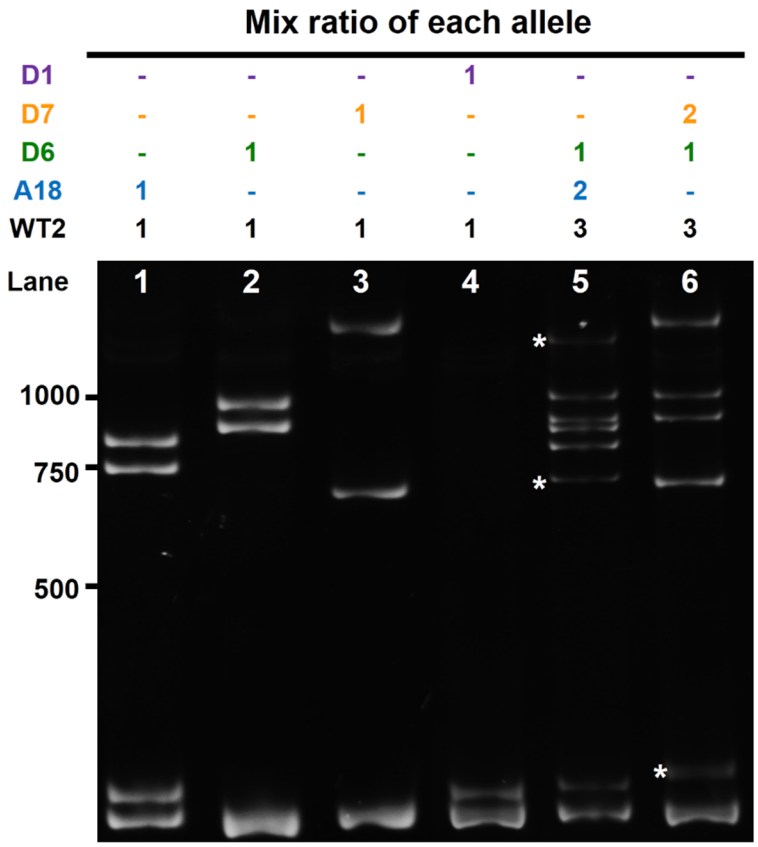


Figure 3-4. PAGE-HMA profile of mixtures of PCR products. PCR products containing only one allele, namely, showing single band in Fig. 2B, were mixed at the same ratio as in Fig. 2B. A heteroduplex formation procedure was performed before PAGE-HMA detection. PAGE-HMA profiles of these mixtures show exactly the same patterns as in Fig. 2B, verifying the patterns as indeed heteroduplexes instead of artificial PCR products. Asterisks highlighted additional heteroduplex bands.

The correlation above can be formulated as follows:

$$B = N^2\text{-}N \tag{1}$$

In Equation (1), N represents the number of different alleles and B the number of Ht bands. I was interested in finding a way to quickly calculate the total number of alleles in a sample by counting the number of its Ht bands. Hence, by defining number of alleles in terms of number of Ht bands, Equation (1) can be transformed into Equation (2) as follows:

$$N = \frac{1}{2}\left(\sqrt{4B + 1} + 1\right) \tag{2}$$

Since the values under consideration must be integers, Equation (2) can be simplified by omitting the small numbers under the square root while imposing a ceiling function to the final result which returns the smallest integer not less than the input:

$$N = \left\lceil \sqrt{B} + \frac{1}{2} \right\rceil \tag{3}$$

Therefore, using Equation (3), the number of alleles can be estimated by counting the number of Ht bands with simple calculation. For example, in Fig. 3-3 B, Lane 7, 8, 12 exhibits 2, 5, and 7 Ht bands, respectively; using Equation (3), we have returns of 2, 3, and 4, which correctly matches the real allele number in each lane. The correctness of Equation (3) has been proven through the simulation of calculating allele numbers from band numbers of 1-200 (See Appendix Fig. 1) using WolframAlpha software from Wolfram Research (Champaign, IL, USA).

The correctness of Equation (3) owes to the ceiling function. Given that N alleles can engender at most N^2-N Ht bands, Ht band number that range from N^2 - N + 1 to $(N + 1)^2$ - (N + 1) should emerge from N+1 alleles. For this reason, the non-integral results from Equation (2) can be rounded up to integer values by the ceiling function.

Note that not all Ht bands can be recognized by PAGE-HMA. In the scenario of Fig. 3-3 D, four alleles are of different abundances, as indicated by the color intensity of the blue header background. Accordingly, their Ht products would show a gradient distribution, as depicted by the intensity of the green background. Some configurations, for example, those framed by the dash line in Fig 3-3 D, might not be seen due to their low abundance. Using Equation (3), we can get the minimal number, but not the exact total number, of alleles in a sample.

In sum, my results proved that PAGE-HMA was capable of simultaneous detection and discrimination of multiple alleles. This is another advantage over TAGE since TAGE converts all Ht configurations into smear bands of T7 digestion product, eliminating the information of Ht variations.

3.2.3 Enrichment of mutant alleles

The detection of Ht by PAGE-HMA can be a strong indicator of a successful GE event. However, final proof comes from sequence analysis of the MT alleles, which is also the most challenging part for embryo microinjection based GE due to the low frequency of the MT alleles among the sample pool. For example, ZFN-mediated GE has been reported to occur in 1~3% cells of zebrafish embryos (Foley et al. 2009; Meng et al. 2008), where only one gene copy is targeted in most cases. Consequently, the targeted allele is usually present at ~1% in a genomic DNA sample of fin clip from a positive individual. Therefore, the validation of successful GE as well as screening for positive individual is a time-consuming task in GE experiment, because it requires the preparation and sequencing of several hundreds of recombinant colonies (Foley et al. 2009; Meng et al. 2008). No effective method was reported to enrich these MT alleles before my report (Chen et al. 2012).

The intactness of PCR products after PAGE-HMA provoked us to develop a procedure for enriching the rare MT alleles. To this end, I developed a procedure to recover the Ht bands from the polyacrylamide gel and then amplify the MT allele through another round of PCR using the recovered Ht as template. I term this procedure gel recovery and PCR (grPCR). To test the effect of grPCR, a DNA mixture of 4 alleles at defined ratio (Column 1, Fig. 3-5 C) was used for first round of PCR and PAGE-HMA. All visible Ht bands were gel-recovered collectively (Dash frame, Fig. 3-5 A) through smashing and soaking with TE buffer. DNA released from these Ht bands was used as templates for subsequence grPCR and PAGE-HMA, where intensity of each band was quantified by densitometry. PCR products from each round were also subjected to TAGE analysis for comparison.

My result demonstrates that this procedure successfully enriched the MT allele. Band densitometry results show that the intensity of dominant WT allele was reduced considerably from 62.0% in the first PCR to 36.0% after one round of grPCR, and was further lessened to 26.3% after another round of grPCR. Meanwhile, intensities of the minority MT alleles increased after first round of grPCR and reached the highest after another round of grPCR (Fig. 3-5 B). This data were gathered through adding up the intensity of bands that belong to the same MT allele. As shown in Fig. 3-6, bands were framed manually for measuring intensity in Gel-Pro Analyzer software. The composition of each band was determined referring to Fig. 3-3. In parallel with this result, in the TAGE profile of the same three rounds of PCR products, the intensity of the digested band increased while the intensity of the undigested WT band decreased after each round of grPCR (Fig 3-5).

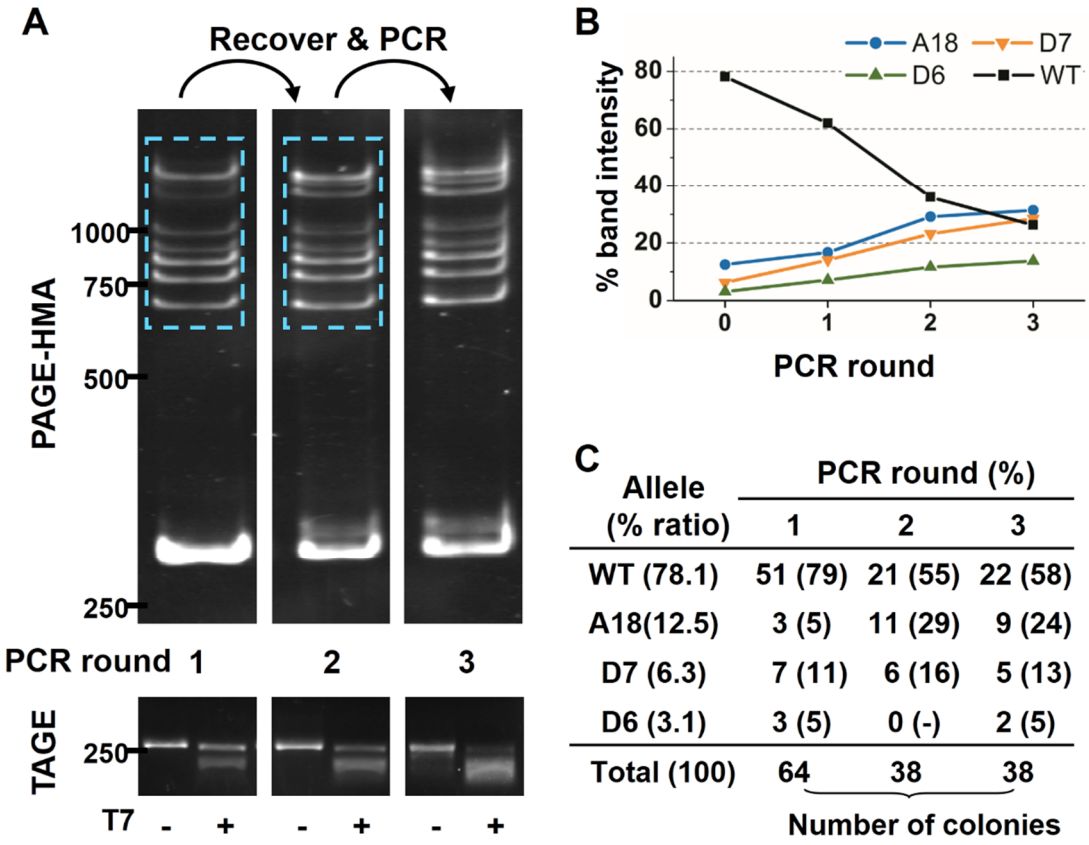
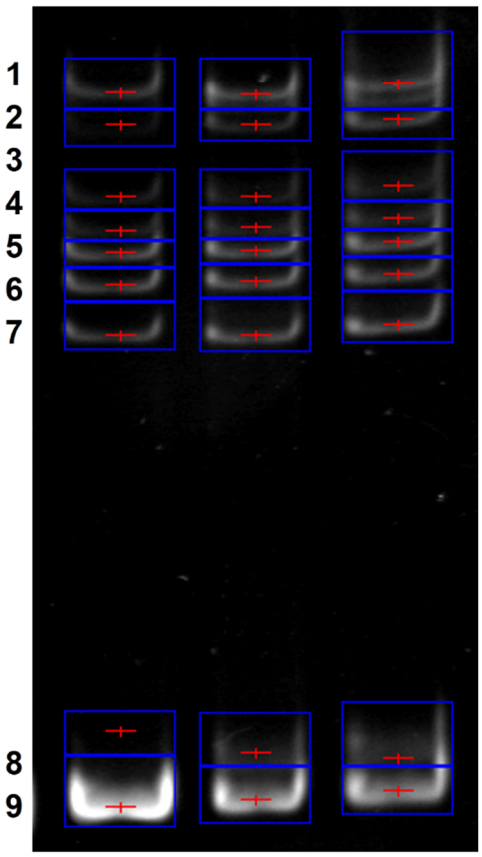


Figure 3-5. PAGE-HMA detection, enrichment and quantification of multiple alleles (In collaboration with X. Zhang, see DNA Res. 2012 Oct;19(5):423-33 & X. Zhang's Thesis)**.** (A) PAGE-HMA and TAGE profiles of multiple alleles after successive rounds of PCR following gel recovery (dash frame). DNA mixture containing 4 alleles shown in lane 12 of Fig. 3-3B was used as template with mix ratios indicated in (C). (B) Relative intensity of different bands of each round. Round 0 represent the initial percentages of each allele. (C) Number of clones and percentages of each allele from PCR products of each round. PCR products of each round was ligated to the vector plasmid for cloning and sequencing.

A



B

Band	PCR round 1	2	3
1	7.3	11.0	17.9
2	1.7	4.3	5.4
3	2.7	4.1	5.5
4	3.6	5.4	5.6
5	7.3	8.7	8.2
6	7.4	9.4	7.7
7	6.8	12.2	10.6
8	1.2	8.9	12.8
9	62.0	36.0	26.3
Sum(%)	100.0	100.0	100.0

C

Allele	PCR round 0	1	2	3
WT	78.1	62.0	36.0	26.3
A18	12.5	16.8	29.2	31.4
D7	6.3	14.1	23.2	28.5
D6	3.1	7.1	11.1	13.8

% band intensity

Figure 3-6. Quantification and enrichment of ZFN-targeted alleles. (A) PAGE-HMA profiles of plasmids mixture shown in Fig. 3. Bands are numbered to the left and framed for quantitative densitometry analysis by the Gel-pro software. (B) Relative intensity/abundance of bands after one, two and three rounds of PCR. (C) Relative intensity/abundance of 4 alleles before PCR (zero) and after one, two and three rounds of PCR, obtained by summing up the corresponding value in (B).

The PCR products were also cloned for sequencing to validate the percentage of each component (Fig. 3-5 C). After one single round of grPCR, the percentage of WT allele decreased from 79% to 55%, and percentage of MT alleles increased relatively. This verifies the enrichment for MT alleles through grPCR.

Taken together, PAGE-HMA profile enables the quantification of each component, while grPCR allows for the enrichment of multiple rare alleles.

It is worth noting that, in the colony sequencing result (Fig. 3-5 C), the last round of grPCR produced no increase of MT alleles. It is because that the composition of the Ht bands remain the same after each round of PCR. Roughly, one half of the Ht band DNA is from WT allele, while the other half is from various MT alleles. Therefore, one round of grPCR is enough for the enrichment of MT alleles. Any additional rounds of grPCR might not increase the proportion of MT alleles, while risking the increase of PCR artifact.

3.3 Procedure optimization

3.3.1 Effect of genomic DNA

To further examine the robustness of my PAGE-HMA method on real GE experiments, I mimic the real situation of detecting indels in a targeted genome by adding medaka genomic DNA to the serially diluted mixtures of the cDNA-originated WT1 and D18 prior to PCR and PAGE-HMA detection. As illustrated in Fig. 3-7, addition of *af* genomic DNA did not interfere with the formation of Ht fragments and their detection by PAGE-HMA. The band pattern and detectable dilution factor remained unchanged with the addition of genomic DNA (compare Fig. 3-2 B and Fig. 3-7).

Nevertheless, detecting indels in genomic context might be disturbed by polymorphisms. As described above, all allelic differences between the primers used,

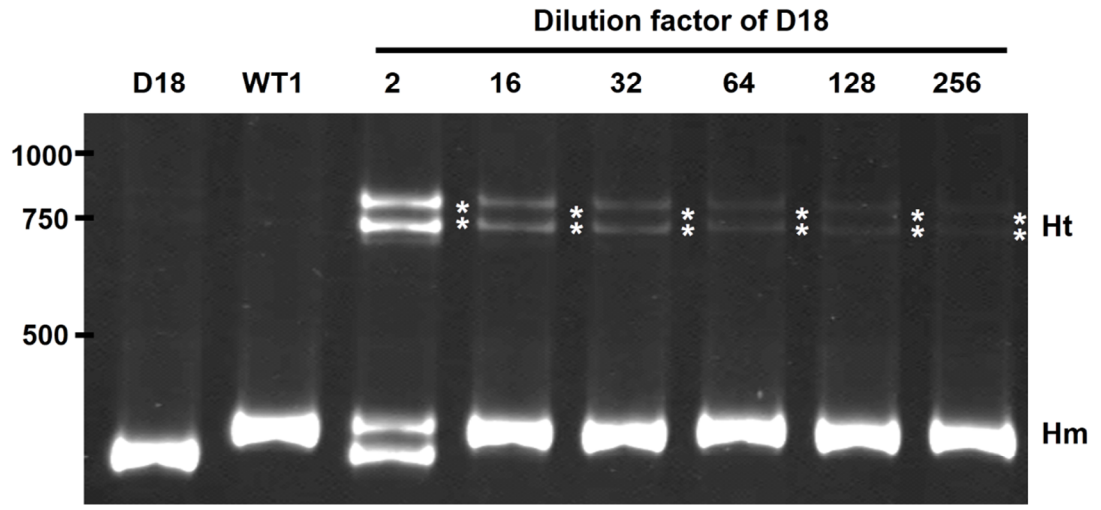


Figure 3-7. PAGE-HMA detection of heteroduplexes produced in the presence of genomic DNA. Templates mixture ratio and PCR conditions were the same as those used for Figure 2B, except that 50 ng of medaka genomic DNA was added to each template. PAGE-HMA profile was not affected by such addition of genomic DNA. Heteroduplexes are highlighted by asterisks. WT1, wild type allele 1; D18, mutant allele with a 18-bp deletion.

either natural polymorphisms or experimentally induced alterations, will give rise to Ht PCR products, which become Ht bands on PAGE. Through gel recovery of the Ht bands (see 3.4.2), I discovered a 10 bp deletion polymorphism existing only at the *gsdf* locus in *af* (Fig. 3-8 B) during the GE experiments, which provides an excellence sample to test the impact of genomic polymorphisms on the PAGE-HMA procedure.

Taking advantage of the 10 bp deletion polymorphism, I examined the PCR products of *gsdf* locus from *af* and *HdrR*, which does not contain polymorphism at the gsdf locus. In addition, GE fish #4 (see 3.6.1) originated from *af* while possessing indels at the *gsdf* locus was included in the experiment. Primer pairs in three lengths were designed: 572, 372 and 305 bp (Fig. 3-8 A). Among these primer pairs, both 572 and 372 include the polymorphism site in their amplicons, while primer pair 305 excludes the polymorphism in its amplicons.

As shown in Fig. 3-8 C, the result of the polymorphism test demonstrates that for medaka line *HdrR*, amplicons from all three primer pairs showed no or very faint Ht bands. Contrarily, for medaka line *af*, Ht bands (Fig. 3-8 C asterisks) from the polymorphism could be seen from amplicons of primer pairs 572 and 372 but not 305. For the *af* originated GE fish #4, which contains various indels at the same *gsdf* locus, Ht bands from both the polymorphism (Fig. 3-8 C asterisks) and the indels (Fig. 3-8 C hash) can be found on the PAGE. Therefore, a WT sample serving as a negative control is required in every batch of the PAGE-HMA detection to rule out polymorphism bands as false positive. This is actually an advantage over the TAGE method, which could not distinguish polymorphisms from indels of interests. In addition, the PAGE-HMA method also provides a tool to survey the GE target site for polymorphism among various populations or strains in prior, so as to determine their suitability for GE experiment.

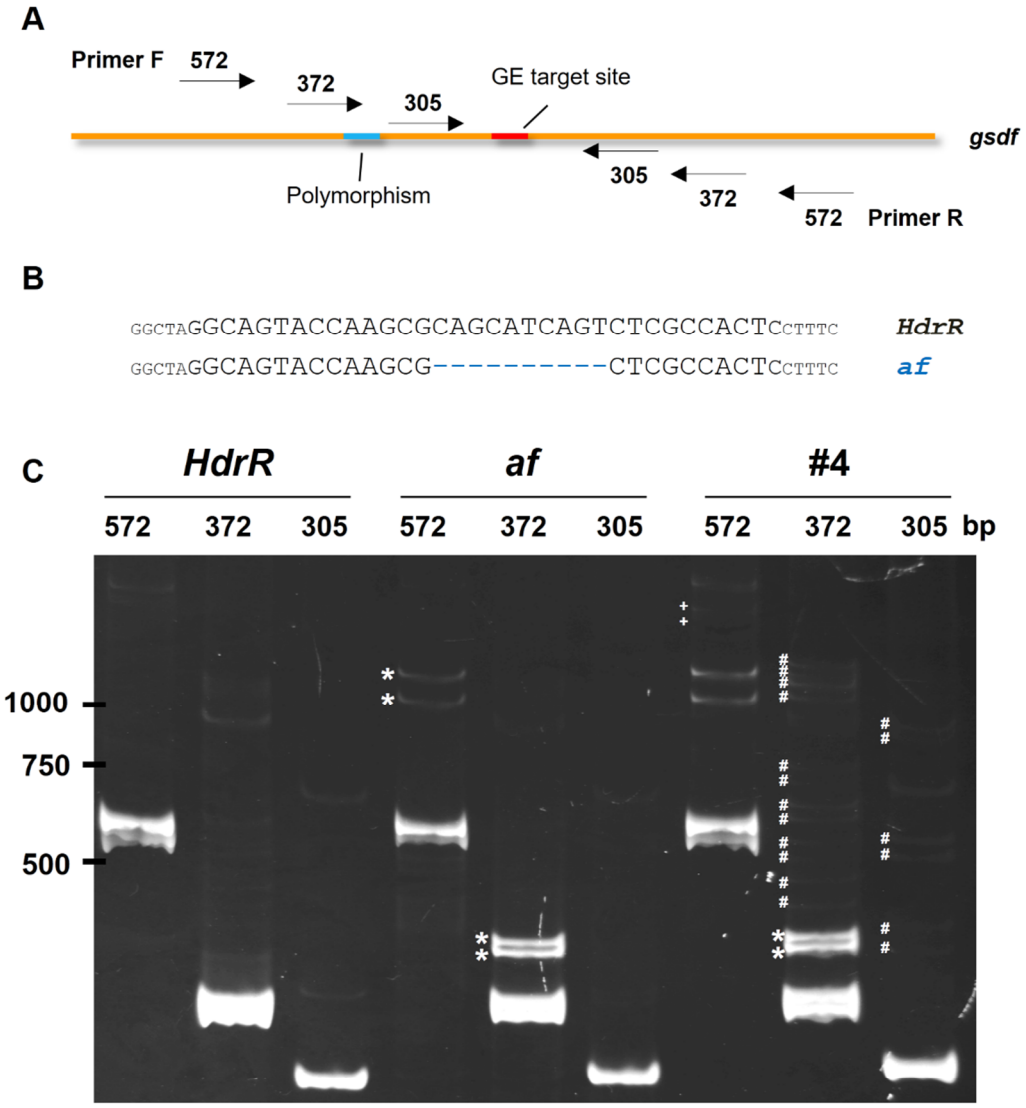


Figure 3-8. Effects of amplicon length and polymorphism. (A) Scheme of primers positions in the genome. Three pairs of primers were used to amplify fragments of 572, 372 and 305 bp with a central position of the disruption site (red). Blue denotes the polymorphism site. (B) Polymorphism found in wild type fish strain *af* as 10 bp deletion. (C) PAGE-HMA profile of PCR products of *gsdf* locus from different strains at various sizes. Natural polymorphism was found by comparing the profile of *af* and *HdrR* at size of 572 and 372 bp. Heteroduplex bands indicating the polymorphism (asterisk) were recovered and sequenced as described and the polymorphism is identified as shown in (B). Using the primer which bypasses this polymorphism site, PCR products showed no heteroduplex bands (305 bp). Thus, to eliminate the impact of naturally existing polymorphism on the PAGE-HMA detection, a wildtype sample is required in every batch of PAGE-HMA. Only samples that show additional bands or different patterns comparing to the wildtype can be deemed as positive. Hash and plus signs indicate such additional bands of the gene disrupted fish #4.

3.3.2 Effect of amplicon length

I would recommend that the length of amplicon for the PAGE-HMA detection should be 200-400 bp. Two reasons attribute to this length. First, the length should be able to include all indels introduced by GE. Usually, these indels vary from 1 to >100 bp (Kim, Kweon, and Kim 2013). The nucleases might cleavage aside the targeting sequence, adding up to hundreds of bp variations to the indel location. Hence, the amplicons should not be smaller than 200 bp to avoid false negative results. Second, consider saving time for electrophoresis, amplicons should not be too large. Smaller DNA can be separated quicker on native PAGE, while larger DNA demand longer electrophoresis. In addition, for the PAGE-HMA method, full separation of Hm and Ht bands is crucial to its sensitivity. For instance, in the polymorphism test (Fig. 3-8), Hm band of 572 bp rested at the middle of the gel while Ht bands of 572 bp remain shrank and hard to identify. As for the separation ability, polyacrylamide gel of 3.3% C is the choice for separation of native DNA. Under the condition of 3.3% C, polyacrylamide gel of 8% T is the best in discriminating DNA between 60-400 bp. Lower T percentage is not recommended for the reason that it might not efficiently separate alleles with minor differences, resulting in smear or merge of bands. To sum up, in my study, the length of amplicons were chosen to be 200-400 bp to guarantee detection of all possible indels without prolonged electrophoresis time.

3.3.3 Detecting homologous indels

The PAGE-HMA detection method takes advantage of the formation of Ht during PCR of mosaic genomic samples which contains both WT and MT alleles. For genomic DNA from F0 individuals, DNA samples should be mosaic. This is because the indel formation efficiency of GE nucleases still below 100%, usually generating alterations in a single allele. Moreover, mRNA of nucleases injected at 1-cell stage

embryos required around 3 h to be translated into functional proteins, while the embryos would have developed to multiple cell stages. The GE events in different cells might vary. Hence, for F0 samples, Ht form naturally during PCR, and no additional step is required.

As for progenies from founder F0 fish, due to the possible existence of homologous fish which contains double MT alleles but no WT allele, detection of Ht requires the additional heteroduplex formation step. In this step, PCR products of the sample are first mixed with WT PCR product at ~1:1 ratio, and then heated to 94°C followed by a slow cool down. Alternatively, the genomic DNA of WT can be added to each sample prior to PCR, so that Ht can be generated during PCR.

3.4 Applying PAGE-HMA to gene editing experiments

The experiments described so far only dealt with model systems by using mixtures of linearized plasmids as templates for PCR. As descripted above, I developed a PAGE-HMA procedure for detection, quantification and enrichment of MT alleles generated from GE experiments. For easy implementation, I pipelined the procedure as shown in the Appendix. I was interested in the development of approaches for direct GE by using engineered endonucleases such as ZFN, TALEN and CRISPR-Cas in medaka. Therefore, in the following subsections, the application of the PAGE-HMA procedure will be demonstrated in real experiments of GE in embryos of medaka fish.

3.4.1 ZFN & TALEN

At first, the ZFN approach was chose for GE in medaka through RNA microinjection into 1-cell stage embryos. The target gene *gsdf* is a gene whose RNA expression is spatially and temporally correlated with early testicular differentiation (Shibata et al. 2010). After GE experiment, several MT alleles were first identified through TAGE, confirming the efficiency of the ZFN approach. Four of these alleles were used in the

multi-allelic model assay described above (Fig. 3-3). For the application of PAGE-HMA, more than hundreds of ZFN-injected embryos were raised into adult fish. Genomic DNA was extracted from fin clips of a subset of samples and subjected to both TAGE and PAGE-HMA. I detected successful GE in 2 out of 48 fish examined. These two fish (#4 and #6) exhibited clear bands of Ht PCR products on PAGE-HMA (Fig. 3-9 A), and a smear or faint band indicative of Ht products on TAGE (Fig. 3-9 B). The band patterns of two fish samples on PAGE-HMA and TAGE were similar to that observed in the previous described model systems (Fig. 3-2 and 3-5). Sequencing 72 recombinant colonies of the Ht PCR products from the third round PCR (a total of 2 rounds of grPCR) revealed 3 different alleles (Fig. 3-9 C). One is the WT allele that was present in 45 clones (63%), the remainder is targeted alleles D7 (22%) and A3D7 (15%). Interestingly, allele A3D7 has both a 3-bp addition and a 7-bp deletion, and more intriguingly, also 4 mismatches of 1~2 bp each at 3 separate positions. It follows that ZFNs can simultaneously introduce addition, deletion and mutation into one and same targeted allele. Hence, ZFN-mediated endogenous GE is possible in medaka embryos, and PAGE-HMA is effective in detecting and cloning ZFN-targeted rare alleles in fish.

Similarly, I further applied the PAGE-HMA method to TALEN-mediated GE experiments performed by my colleague. As shown in Fig. 3-10, the PAGE-HMA method also detects the MT of a 3 bp addition generated by TALEN.

3.4.2 CRISPR-Cas

The CRISPR-Cas system has recently emerged as an efficient platform for gene editing and is becoming a routine method for the study of gene function due to its simple and robust setup. To test the feasibility of the CRISPR-Cas system on medaka and to expand the application of the PAGE-HMA method, I utilize the CRISPR-Cas

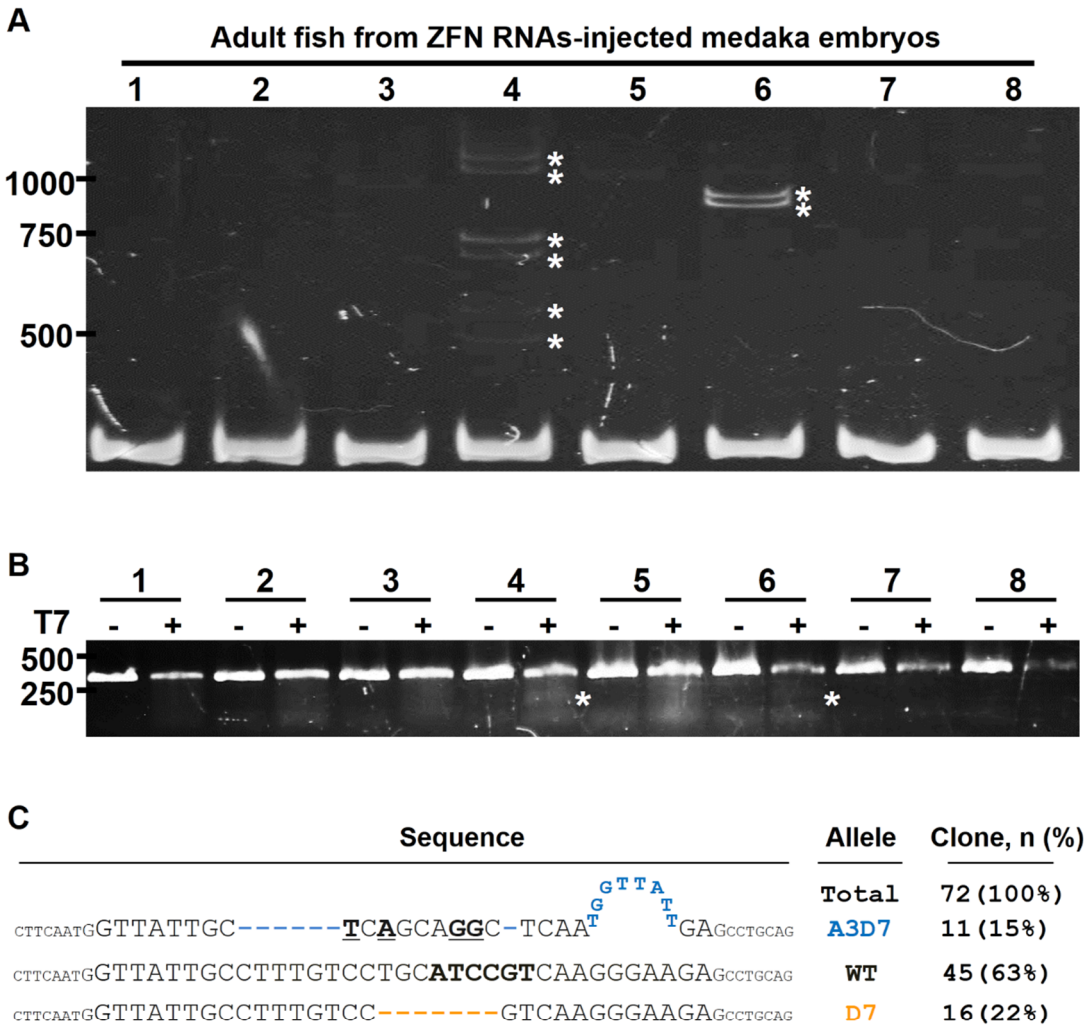


Figure 3-9. Detection of ZFN-mediated allelic alterations in medaka (In collaboration with X. Zhang and G. Guan, see DNA Res. 2012 Oct;19(5):423-33 & X. Zhang's Thesis)**.** Medaka embryos were microinjected with RNAs for a pair of ZFNs that targeted the first exon of *gsdf* locus and were raised into adult. Genomic DNA was extracted from fin clips of eight randomly sampled adults and was subjected to PAGE-HMA and TAGE detection for genome editing events. (A) PAGE-HMA profiles of PCR products, clearly showing the presence of heteroduplex DNA (asterisks) in two out of eight fish. (B) TAGE profile of the PCR products. Faint bands of heteroduplex DNA (asterisks) within smears can be seen from the same two fish. (C) ZFN-disrupted alleles in medaka sample #4. Three different alleles of the medaka *gsdf* gene from sample #4 are compared by partial sequence alignment. Additions are shown as a loop. Deletions are depicted by gaps. Mismatches are underlined.

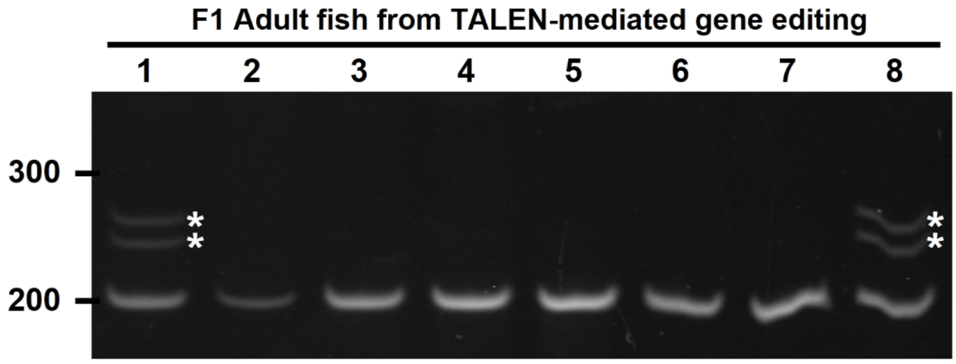


Figure 3-10. Detection of TALEN-mediated genome editing in medaka (In collaboration with T. Wang)**.** Fin clips from F1 medaka adults of TALEN-mediated *dnd* gene knockout experiments were examined by PAGE-HMA. Asterisks highlight heteroduplexes.

system from Chen and Wente lab (Jao, Wente, and Chen 2013, 201) for GE in medaka embryos. The gene *nanog,* which regulates proliferation during early fish development (Camp et al. 2009; Sánchez-Sánchez et al. 2010; Schuff et al. 2012), was selected for GE. Two target sites at the first exon of nanog were designed with the corresponding gRNA vectors constructed and named as gRNA 183 and gRNA 319, respectively, according to their position in the genome. Cas9 RNA and gRNA were transcribed and microinjected to 1-cell stage embryos of medaka *HdrR II* strain. The toxicity of this CRISPR-Cas system was tested first with 154 embryos co-injected with ~4 μL of 100 ng/μL gRNA 319 and 100 ng/μL Cas9 mRNA. All injected embryos hatched successfully and survived the first 3 days, suggesting a low toxicity of this system. Then, for a quick test of the efficiency of the CRISPR-Cas9 system, genomic DNA of the 3dpf microinjected embryos from both gRNA 183 and 319 were extracted by pooling two embryos in one sample. A total of 12 embryos from each gRNA were examined using the PAGE-HMA. As shown in Fig. 3-11 B, Ht bands can be seen in almost all lanes, suggesting the success of GE. Some of these Ht bands (Dash frame, Fig. 3-11 B) were recovered for cloning and sequencing. The sequencing result demonstrates the successful GE events in these embryos, confirming that the CRISPR-Cas system can work well in medaka fish.

Note that gRNA 319, with shorter targeting site, generated more Ht bands compared with gRNA 183. According to the guideline for designing a gRNA, the targeting sequence should be at least 23 bp (Jao, Wente, and Chen 2013). However, my result suggests that gRNA targeting sequence as short as 21 bp might also work with sufficient efficiency.

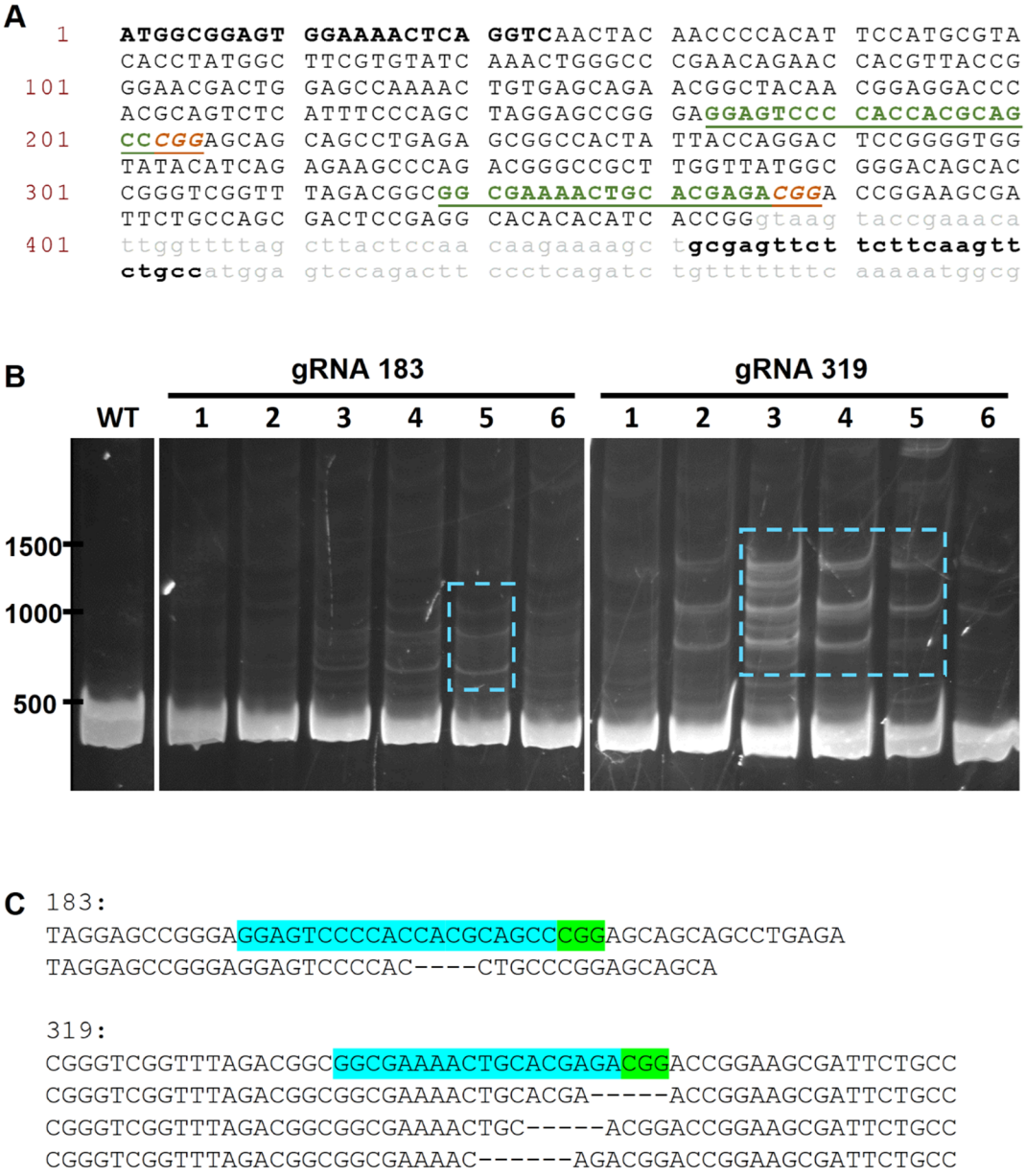


Figure 3-11. Detection of CRISPR-Cas mediated allelic alterations in medaka. Two gRNAs, gRNA 183 and 319, were used for CRISPR-Cas experiment. Genomic DNA of 3 dpf embryos was extracted by pooling 2 embryos in each sample. PCR was performed and amplicons were analyzed by PAGE-HMA. (A) Targeting site of gRNA 183 and 319 (name according to position). (B) Heteroduplex bands indicating successful genome editing events. (C) Sequencing results of the recovered bands as framed in (B)

CHAPTER 4 Results and discussion II:

RNA visualization by morpholino MB (MOMB) *in vivo*

As reviewed above, morpholino oligos (MOs) have been widely used in developmental biology studies due to its comparatively low toxicity, high stability and specificity *in vivo*. These traits encouraged me to combine the use of MOs with the MB design to provide a new solution to the visualization of RNA dynamics. Therefore, a new type of molecular beacon (MB) with morpholino backbone was designed and termed MOMB for convenience. MOMB might overcome some shortcomings of the current MB systems such as instability, the specificity and so on.

To determine whether the MOMB can work as expected, *in vitro* assays were performed. The results show that MOMB has good potential for *in vivo* RNA tracking. Hence, a series of *in vivo* assay were carried out in medaka embryos to examine the toxicity, stability and specificity of MOMB. The results of these assays will be presented and discussed in the following sections.

4.1 MB design and preparation

4.1.1 Design of MBs

For a successful design of MOMB, the thermodynamics is the most crucial factor to be considered. Conventional MBs have Tm of 50-60 °C for the stem and Tm of 70-80 °C for the duplex of MB and target as explained in section 1.3.3.1, which provides a good starting point for the design of MOMB.

As for the design of the stem, the Tm of MO homoduplexes is to be considered. Only one previous report (Ouyang et al., 2009) proposed a calculation for Tm of MO homoduplexes based on the Tm of a set of MO duplexes linked by a dimethoxynitrobenzyl linker, measured by thermal denaturation curves:

$$Tm = 1.9 \times (A + T) + 5.7 \times (G + C) \quad (4)$$

The results calculated from Equation 4 matched their observation well. Thus the Equation 4 was used to guide my design of MOMB, which suggests a stem of 8-11 bp C:G pairs to achieve a Tm of 50-60 °C for the MOMB itself. On this basis, to further understand the effect of different lengths, three MOMBs were designed with stem lengths of 5, 7 and 10, respectively (Fig. 4-1). MOMB-5 and MOMB-10 were targeting medaka *nanog* and *rx2* mRNA respectively, while MOMB-7 used random sequence without known target mRNA. The stem-loop structures of the three MOMBs were simulated and confirmed using Quikfold application from the DINAMelt Web Server (Markham and Zuker 2005; Markham and Zuker 2008).

As for the design of the recognition sequence, I referred to several guidelines for the MOs provided by the manufacturer. The three MOMBs were all designed according to these guidelines: a) total length no more than 25 nt, b) GC content of 40 - 60% with less than 10 guanine residues, and c) no contiguous stretch of four or more guanine residues. Also note that the manufacturer only provides MOs with DABCYL (4-([4-(dimethylamino)phenyl]azo)benzoic) as quencher, and with fluorophore modifications at 3' end. To guarantee a good water solubility, fluorescein was chosen as the fluorophore. To make full use of the 25 nt length limitation of the MOMB to reach higher Tm of the MOMB-target duplex, shared stem strategy was adopted (Tsourkas, Behlke, and Bao 2002a). The shared stem which is complementary to the target sequence was set at the quencher end, so that the fluorophore end is

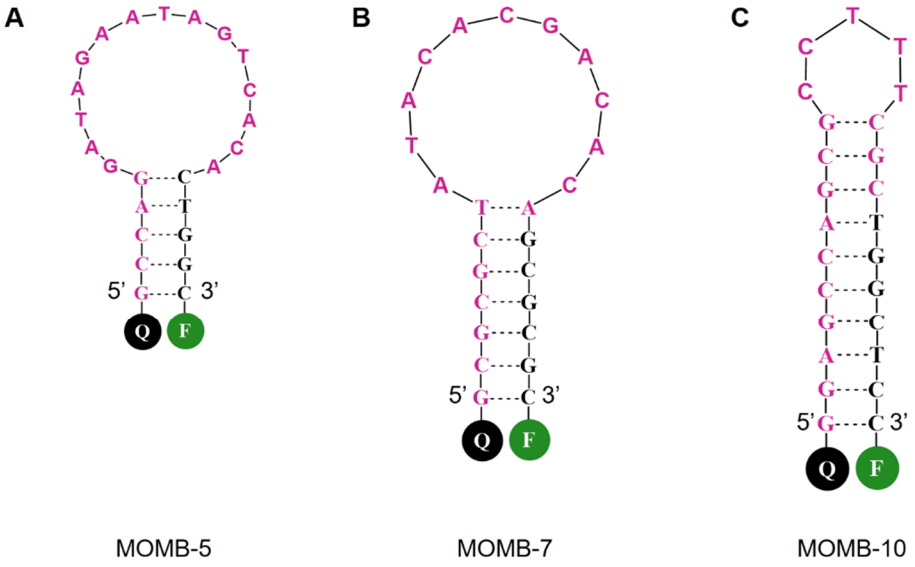


Figure 4-1. MOMB designed with various stem length. Three MOMBs were designed with different stem lengths and mRNA targeting sequences. (A) MOMB with 5 bp stem targeting medaka *nanog*. (B) MOMB with 7 bp stem without target gene. (C) MOMB with 10 bp stem targeting medaka *rx2*. Q, quencher; F, fluorophore.

dissociative to exempt the quenching effect from guanines of the target strand (Medley and Zhu 2013; Seidel, Schulz, and Sauer 1996).

4.1.2 Preparation and purification of MOMBs

MOs were provided as dry powder. Upon receiving the MOMBs, they were dissolved in ultrapure water as 1 mM stock solution. Thanks to the high solubility of the morpholino, the powder dissolved rapidly in pure water, as opposed to another neutrally charged backbone PNA, which might require the addition of organic solvents, and could hardly reach the concentration of 1 mM as a convenient stock solution.

In the preliminary test, the MOMBs alone exhibited very high background fluorescence signals. These MOMBs should show very low fluorescence with the stem-loop conformation when unbound with their target. According to the manufacturer, selective precipitation was used for purification of the final product, which might leave certain impurities in the product such as dissociative fluorophores, or incomplete MOs of shorter length and without the quencher. To remove these impurities, a centrifuge ultra-filter of 3kDa cutoff was used. Since the molecular weight of MOMBs are larger than 9kDa, they would remain above the ultra-filter membrane after centrifugation, while the smaller impurities would go through the membrane and be discard. As expected, after the purification process, the background fluorescent signal of the unbound MOMBs was reduced significantly while their signal-to-noise (S/N) ratios (with target/without target) increased relatively. Therefore, all MOMBs went through this purification process before use.

4.1.3 Stem length effect

Without knowing the optimal stem length for MOMB, three MOMBs were designed with stem length of 5, 7 and 10 as shown in Fig. 4-1. To select the optimal length, the

Tm of each MOMB were first determined from their thermal denaturation profile as described in section 2.4.2.2. As shown in Fig. 4-2, MOMB-5 has a Tm of less than 55 °C, while MOMB-7 has a Tm of 60 °C and MOMB-10 has a Tm of about 70 °C. Similar to conventional MBs, MOMB with a longer stem show a higher Tm. Note that using Equation 4, the Tm for each MOMB should be 24.7, 36.1 and 49.4 °C, respectively, which is much lower than the observed values. This could be explained by the difference of the linkers between the homoduplexes. As mentioned above, Equation 4 was used for MOs linked by a long dimethoxynitrobenzyl-based linker (Ouyang et al. 2009). Such a long and flexible linker could compromise the stability of the MO duplex in that it increases the possibility of different conformations other than the desired stem loop. Contrarily, the MO duplex in MOMBs are linked by short MO sequences, which has less possibility for different conformations.

From the results above, all MOMBs seem to have proper Tm values. Therefore, I examined their S/N ratio in phosphate buffer (PB: 1 mM Na_2HPO_4 + KH_2PO_4 for pH=7, 100 mM NaCl, 3 mM $MgCl_2$) with 5 times of perfect match (PM) target. Surprisingly, MOMB-7 show the highest S/N ratio of more than 12, while both the MOMB-5 and MOMB-7 are only around 2. This result indicates that 7 bp is the optimal stem length among the tested three. Therefore, the MOMB-7 was chosen for the subsequent experiments and is term directly as MOMB.

4.2 Characterization of MOMB *in vitro*

To verify that MOMB is correctly synthesized and purified, certain characterization assays were performed. In addition, before using MOMB for *in vivo* assays, its potential for *in vivo* applications were determined *in vitro*. RNA probes for *in vivo* application should possess certain features, including salt-independent hybridization, resistance to nucleases and binding proteins, and high specificity.

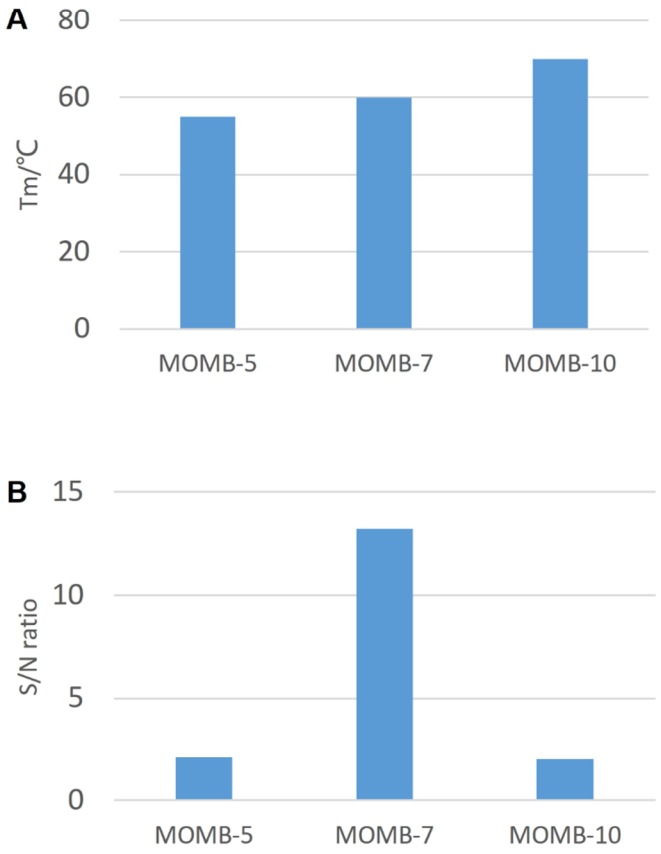


Figure 4-2. Tm and S/N ratio of MOMB-5, -7, -10. (A) Tm value of each MOMB obtained from their respective thermal denaturation curves. (B) The S/N ratio of each MOMB with their respective perfect match target.

4.2.1 Mass spectroscope and UV-Vis absorption spectra of MOMB

Before testing the potential of the MOMB (MOMB-7) for *in vivo* application, I examined the basic characteristics of the molecules to ascertain that they are correctly synthesized and purified. Firstly, MALDI-TOF/TOF mass spectroscope profile of the MOMB was obtained from Gene Tools as quality control document. As shown in Fig. 4-3 A, the mass spectrum of MOMB show a main peak at 9428 Da, which matches the theoretic molecular weight of 9430. Also note that the mass spectrum contains another peak at 4006, which could be the doubly charged morpholino molecules (Hudziak et al. 1996).

Next, the UV-Vis absorption spectra of the MOMB was compared with that of DNA and MO (Fig. 4-3 B). As an analog of DNA, MO has an absorbance peak at 265 nm, close to that at 260 nm of DNA. MOMB also show an obvious peak at 265 nm, corresponding to its morpholino backbone. Moreover, a smaller wider peak at around 490 nm can be seen from MOMB, which denotes the modification of fluorescein and DABCYL. These results confirm the successful synthesis and purification of the MOMB, allowing further examination on the behaviors of MOMB.

4.2.2 Thermal denaturation profiles of MOMB

To confirm that MOMB possessed the desired characteristics of conventional MBs, the thermal denaturation profiles of MOMB was investigated. Such profiles were plotted by monitoring the fluorescence change from 20 to 96 °C, through which a sigmoidal curve can be seen for conventional MBs as shown in Fig. 4-4 A (Vet and Marras 2005). In the absence of the target, the MB shows a low fluorescence at low temperature due to the stem-loop structure. As the temperature increases, the stem deplexes denature and give way to the random-coil configuration, emitting increased fluorescence. Contrarily, in the presence of the target, the MB shows a high

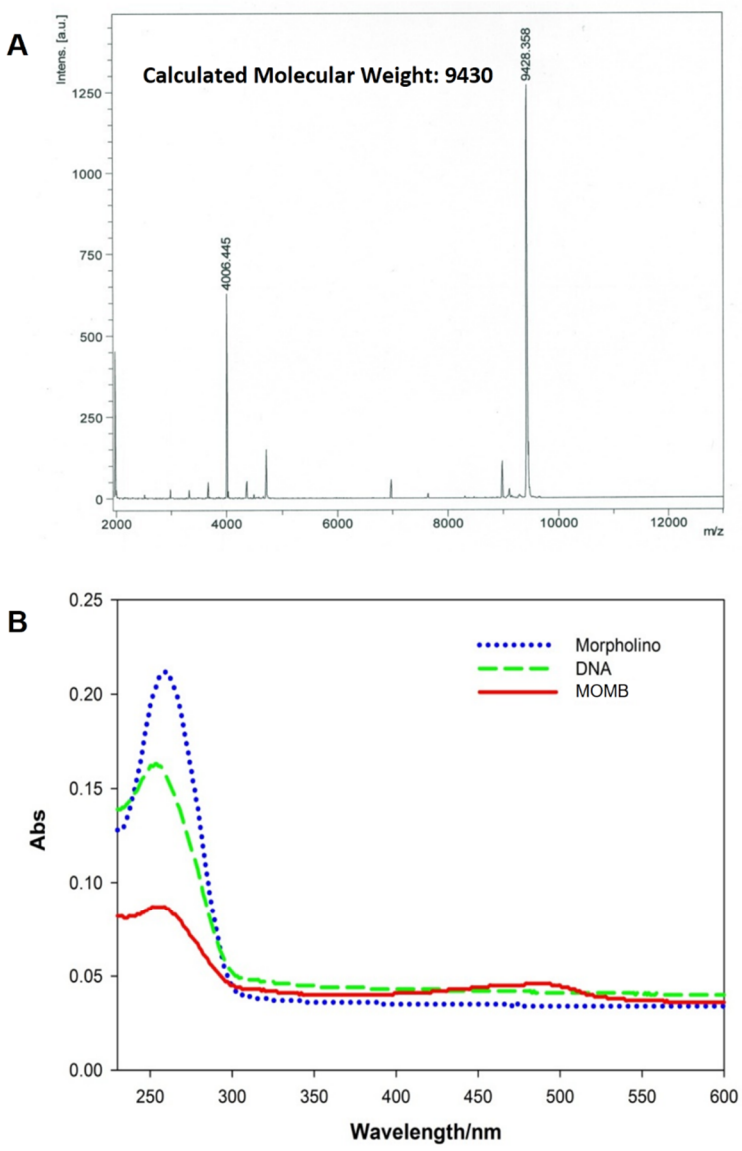


Figure 4-3. MALDI-TOF/TOF mass spectroscope and UV-Vis absorption spectra of MO-MB. (A) MALDI-TOF mass spectroscope profile of MOMB-7 (Provided by Gene Tools company) (B) UV-Vis absorption spectra of linear morpholino, DNA and MOMB.

fluorescence at low temperature, but the signal diminishes gradually as the temperature increases. It is because the probe-target duplexes become destabilized and the stem-loop structure is restored. However, as the temperature increases further, the stem-loop structure also become denatured, resulting in a slight increase of the fluorescence.

The thermal denaturation profiles of MOMB and MOMB-target duplex were obtained in PB as shown in Fig. 4-4 B. The two curves show the same sigmoidal profiles akin to that of conventional MBs (S. Tyagi and Kramer 1996; Vet and Marras 2005), and can be explained likewise (Bonnet et al. 1999; Vet and Marras 2005). Such profiles confirmed that MOMB forms stem-loop conformation and probe-target duplex as desired.

4.2.3 Salt dependence of RNA binding

Conventional MBs require certain salt concentrations to stabilize their stem-loop structures as well as the probe-target duplexes. Nevertheless, cytoplasm provides only low salt environment (Pogorelov et al. 2006; Slack, Warner, and Warren 1973), which hinders the function of conventional MBs. To verify the salt-independent hybridization ability of MOMB, its thermal denaturation profiles were obtained in different buffer. First, the MOMB alone was examined in pure water to check whether the stem-loop structure remains. As shown in Fig.4-4, MOMB show a sigmoidal profile similar to that of conventional MBs (Vet and Marras 2005), indicating that MOMB could maintain a stable stem-loop conformation in the absence of ions.

I further examined the effect of different ion strength on the recognition of target sequence by MOMB using DNAMB as a comparison. The thermal denaturation profiles of MOMB-target duplexes were obtained as shown in Fig. 4-5 B. The Tm value of MOMB remain the same in the buffer series with concentration of NaCl from

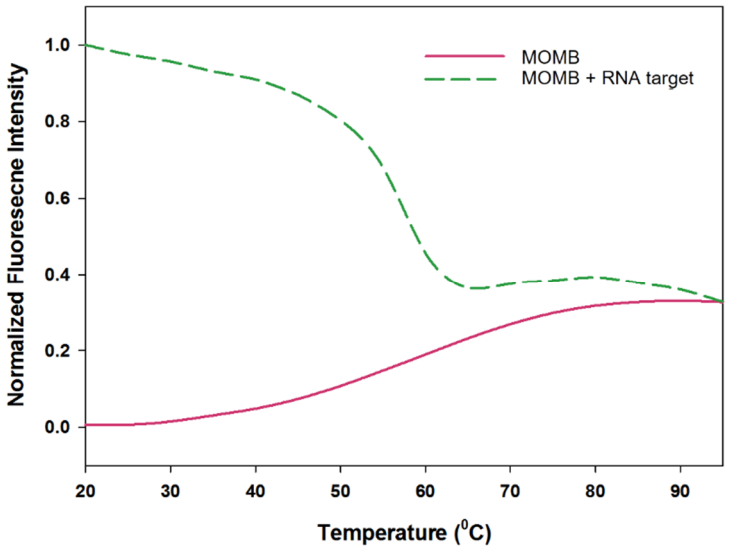


Figure 4-4. Thermal denaturation profile of MOMB. (A) Theoretic thermal denaturation profile of conventional MB. Adpated from Vet and Marras, 2005. (B) Thermal denaturation profile of MOMB and MOMB-target duplexes.

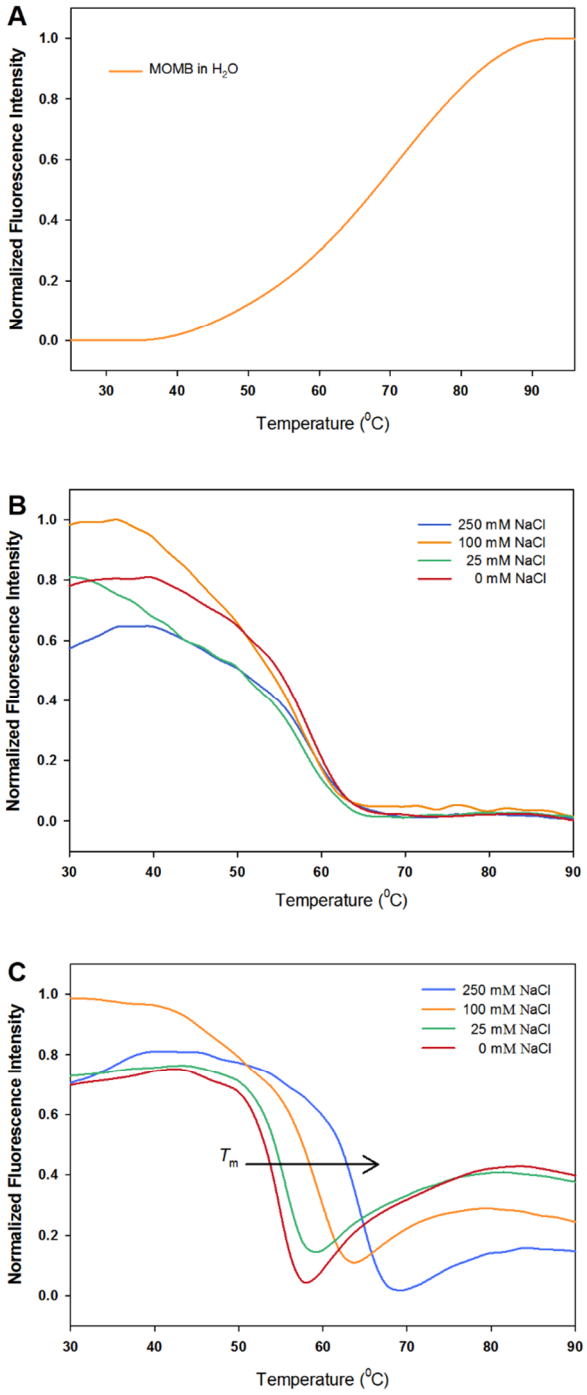


Figure 4-5. Thermal denaturation profiles of MOMB and DNAMB in different salt concentrations. (A) Thermal denaturation profiles of MOMB in water. MOMB showed the same sigmoidal profile as in buffer with the same tm. (B) Thermal denaturation profiles of MOMB-target duplexes under various ion strengths. The Tm of MOMB remained the same. (C) Thermal denaturation profiles of DNAMB-target duplexes under various ion strengths. The Tm of DNAMB increased as the ion strength increase.

0 mM to 250 mM. In contrast, the Tm of DNAMB was affected by ion strength (Fig. 4-5 C). In pure H_2O, the Tm of DNAMB was only 54 °C, which is 11 °C lower than that in 250 mM NaCl. This result verified the unique advantage of MOMB for *in vivo* applications. Thanks to the nonionic morpholino backbone of MOMB, the low salt environment does not compromise the function of MOMB (Tercero et al. 2009; Zu et al. 2011). Contrarily, for conventional DNAMB or other backbones such as LNA and 2-OMe RNA, low ion strength might be unfavorable since their stem-loop and MB-target stability require higher ion strength (Kuhn et al. 2002).

4.2.4 Resistance to nuclease digestion

For conventional MBs, a major hurdle for their application in living organisms is that they could be digested by various nucleases (J. J. Li, Geyer, and Tan 2000; Molenaar et al. 2001; Tsourkas, Behlke, and Bao 2002b). These nucleases include RNase H which could cleavage DNA:RNA duplexes, as well as DNase and RNase which non-specifically degrades DNA and RNA, respectively. As artificial molecules, MOs have been reported to be insensitive to most nucleases (Hudziak et al. 1996). To further confirm this, MOMB and DNAMB were both subject to nucleases disruption assays.

First, MOMB and DNAMB were tested against DNase I, which is a nonspecific DNA endonuclease. As seen in Fig. 4-6 A, after the addition of DNase I, the fluorescent signal of DNAMB increased gradually and reached an equilibrium after 75 min, indicating the thorough release of fluorophores due to the digestion of DNAMB. In contrast, MOMB was not affected by the DNase even after overnight incubation.

Then, MOMB and DNAMB were tested against RNase H. RNase H is an endoribonuclease that specifically degrade RNA which is hybridized to DNA. RNase H is so effective in cleaving RNA:DNA duplexes that it has been utilized for RNase H dependent gene knock-down (Summerton 1999). In my RNase H assay, MOMB and

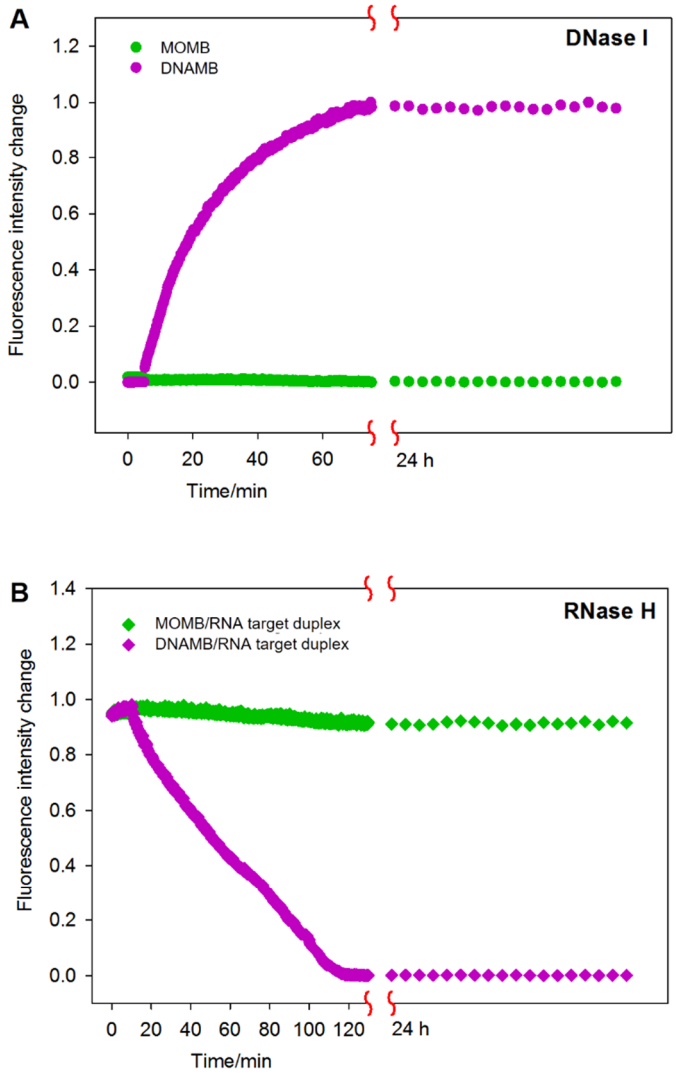


Figure 4-6. Enzyme resistance of MOMB and DNAMB against DNase I and RNase H. (A) Fluorescent dynamics of MOMB and DNAMB with addition of DNase I. (B) Fluorescent dynamics of MOMB and DNAMB with addition of Rnase H.

DNAMB were hybridized with PM RNA target to form DNAMB:RNA or MOMB:RNA duplexes, followed by the addition of RNase H. As demonstrated in Fig. 4-6 B, the DNAMB:RNA duplexes were gradually digested by RNase H, which resulted in the decrease of RNA target as well as the fluorescent signal emitted by the open conformation of DNAMB. As for MOMB, the fluorescent signal did not change even after overnight incubation, which was also in concordance with the previous reports that MO:RNA duplexes are insensitive to RNase H (Summerton 1999).

4.2.5 Resistance to protein binding

Aside from nucleases, certain DNA/RNA binding proteins in the intracellular environment can also open the stem-loop conformation of MBs, which become another major challenge for the application of MBs in living cells (Boutorine et al. 2013). Single strand DNA binding proteins (SSB) is one of the representative protein that has proved to disrupt the stem-loop of DNAMB (J. J. Li et al. 2000). Whether MOs can be affected by SSB is not reported. Therefore, effect of SSB on MOMB and DNAMB were examined. As depicted in Fig. 4-7, DNAMB showed 5-fold increase of fluorescent signal after the addition of SSB, indicating the disruption of the stem-loop conformation. The fluorescent signal of MOMB remained the same after SSB addition, demonstrating that MOMB cannot be affected by SSB. This might largely due to the nonionic and the 6-member ring backbone of morpholino which is not compatible to the binding proteins for negatively charged DNA/RNA with 5-member ring backbones.

4.2.6 Sequence discrimination ability of MOMB

Finally, the hybridization of MOMB and target was studied. For *in vivo* application, the MOMB should show quick response to target RNA to achieve real-time visualization. Moreover, MOMB should avoid false positive signal due to the binding

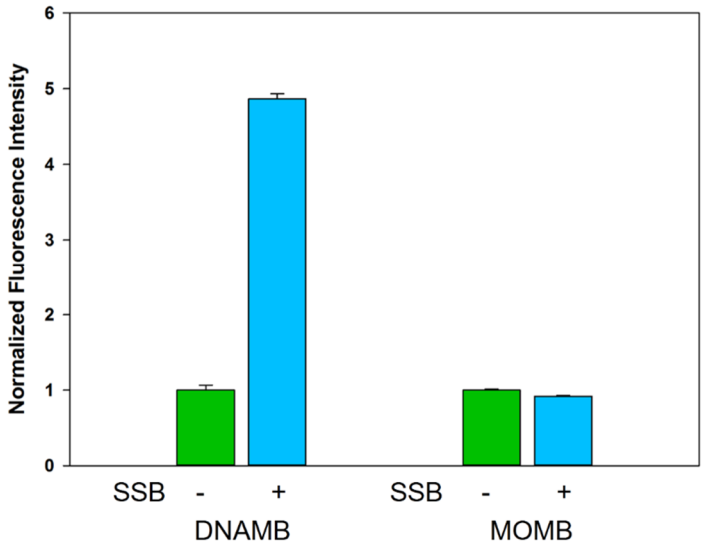


Figure 4-7. SSB effect on DNAMB and MOMB. Fluorescent intensity of DNAMB and MOMB before and after addition of SSB.

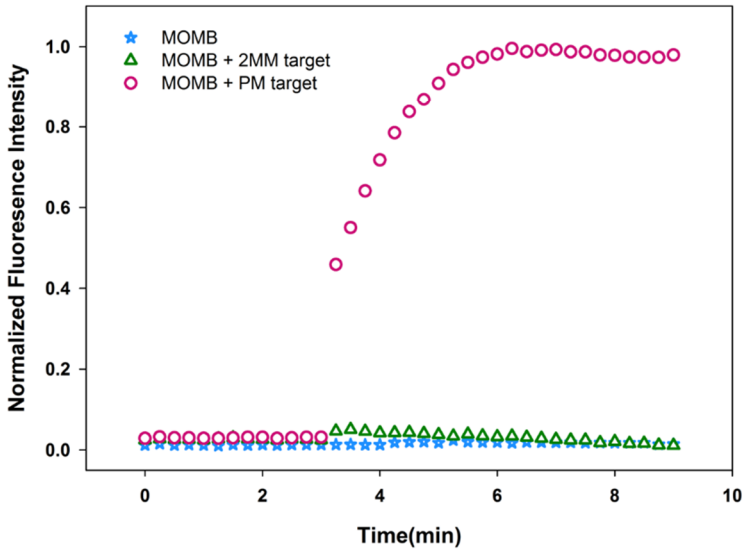


Figure 4-8. MOMB hybridization kinetics. Fluorescence kinetics of MOMB after the addition of perfect match (PM) target and 2 nucleotides mismatch (2MM) target.

of non-specific RNA. To this end, the fluorescence kinetics of MOMB were monitored to access the response speed and the specificity of MOMB. As shown in Fig. 4-8, after the addition of perfect match target (PM) of MOMB, the fluorescent signal increased rapidly and reach equilibrium within only 3 min, demonstrating the quick hybridization of the MOMB to its target. In comparison, the addition of target with 2 nt mismatch (2MM) had no effect on the fluorescent signal, verifying the specificity of MOMB.

In conclusion, the *in vitro* assays demonstrated that MOMB possessed all basic characteristics of conventional MB. In addition, it could tolerate low ion strength, the digestion of nucleases, and the binding of SSB. Also, it demonstrated a high specificity with the ability to discriminate targets with 2 nt mismatch. Encouraged by these results of *in vitro* assays, I carried on to determine the feasibility of MOMB for RNA monitoring in living embryos.

4.3 Characterization of MOMB *in vivo*

Almost all live-cell assays with RNA probes were performed in cell culture system, and only a few were reported in acute brain slice of mouse (Park et al. 2014) or unfertilized oocytes of *drosophila* (Bratu et al. 2003). Comparing with these live-cell system, developing embryo is a far more dynamic system with dramatic gene expression and morphology change (Arbeitman et al. 2002; Tadros and Lipshitz 2009), which could be problematic for RNA visualization. To date, no probe has been reported for real-time RNA visualization in the developing embryo.

Three major aspects are to be concerned for *in vivo* RNA visualization: a) the biocompatibility, or the toxicity, of the probe to the development of the embryo, b) the long-term stability of the probe, and c) the specificity of the probe. In the following sections, these three aspects of the MOMB is investigated and discussed.

4.3.1 Biocompatibility of MOMB *in vivo*

The first character examined *in vivo* is the biocompatibility of MOMB. For effective RNA tracking *in vivo*, the probe used should not affect the normal biological process of the organisms, which is equivalent to a good biocompatibility. The developing embryos can serve as a best model for testing the biocompatibility, in that the developing process involving thousands of genes is highly sensitive to foreign molecules. To this end, 100 μM MOMB and DNAMB were injected to1-cell stage medaka embryos, respectively, and the development process of these embryos were monitored during the first 24 h. As shown in Fig. 4-9, embryos with MOMB injection show no different in morphology comparing with the WT uninjected embryos, while embryos with DNAMB injection exhibited sever developmental delay and deform in the first 6 h after injection and finally ended up with death after 24 h, leaving no recognizable cell structure or fluorescent signal from the DNAMB. Though DNAMB were reported to be biocompatible via experiments in cultured cells, it showed significant toxicity to early developing embryos. This result confirmed the good biocompatibility of MOMB to developing embryos.

4.3.2 Toxicity assessment during embryonic development process

With the good biocompatibility result, I further investigated the toxicity of MOMB throughout the embryonic development of medaka in a statistic manner. The experimental design was based on the Fish Embryo Acute Toxicity Test form OECD guideline (OECD 2013). Two doses of MOMB, 50 and 250 μM were injected along with 5 mg/mL Texas Red dextran to ~150 1-cell stage embryos, respectively. The non-toxic Texas Red can server as a reference dye to rule out unsuccessful injection. To determine the toxicity, the morphology of embryos was assessed at three time points: 1 dpf, 5 dpf and 3 dph. Such a time point setting was to evaluate the effect of MOMB

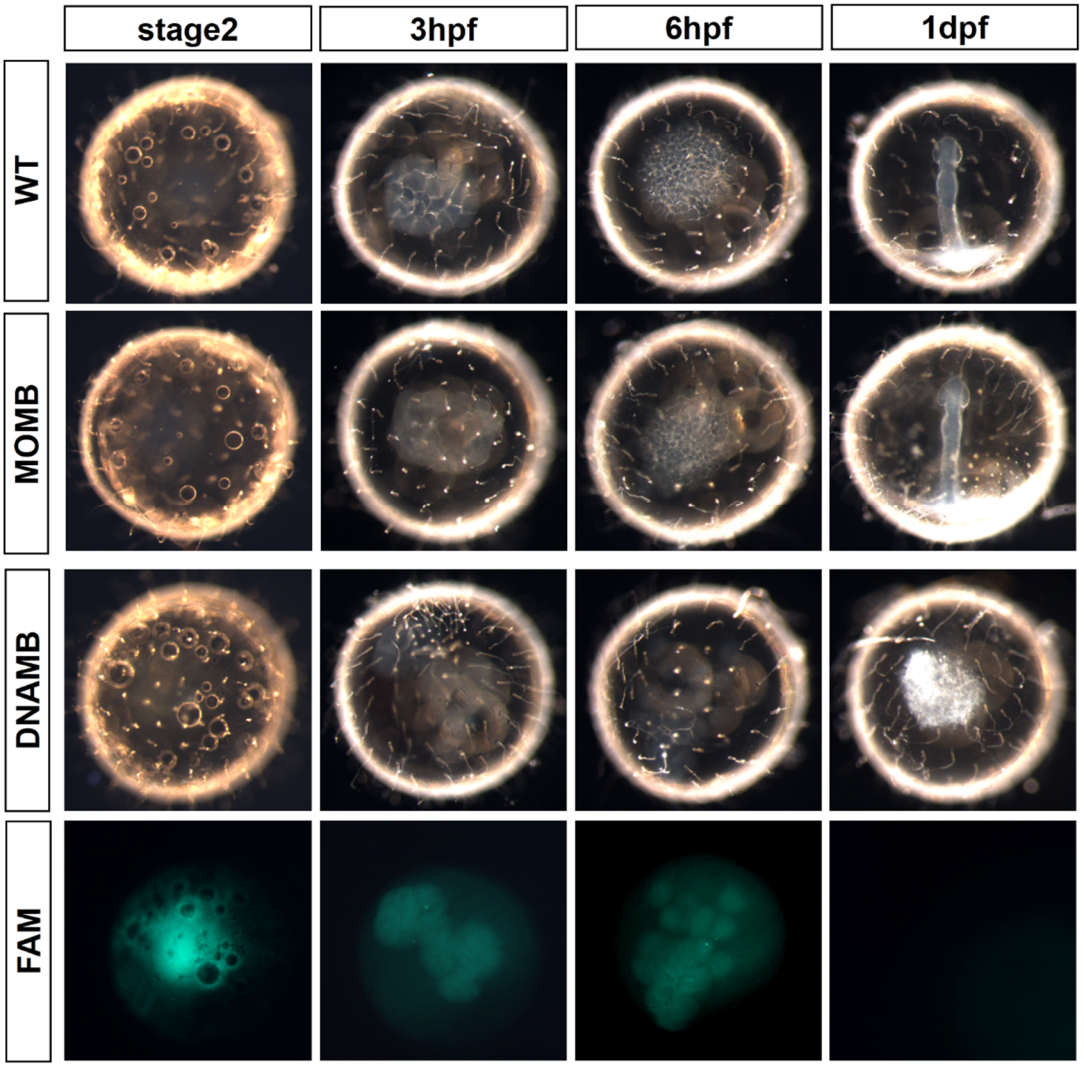


Figure 4-9. Biocompatibility of MOMB and DNAMB. 100 μM of MOMB or DNAMB was injected to 1-cell stage embryos to preliminarily evaluate their biocompatibility.

on the following development processes: a) the gastrulation and the epiboly which complete 1 dpf on stage 28, b) somite, head and cardiovascular system development of 5 dpf embryos on stage 34, and c) hatching and swimming of fry on 3 dpf (Iwamatsu 2004). Each embryo was observed and categorized into one of the three groups: normal, abnormal and dead. For fry on 3 dph, not hatching was considered dead and failure to swim is considered abnormal.

The proportion of the embryos for the toxicity assay at each time point is summarized in the bar chart of Fig. 4-10. Typical embryos of each time point are also demonstrated in Fig. 4-10, and representative embryos of each category can be found in the Appendix Fig. 1. The result of this assay showed a very low toxicity of MOMB to medaka embryos. The percentage of normal embryos in both MOMB injection groups were comparable to that in the Texas Red injected or WT uninjected group at every time point examined (see Appendix Table 3 for detail). Even at the significant high injection concentration of 250 μM, the survival rate of the 3 dph fry could reach 84.5%, close to the 89.6% of the WT group. Though it was slightly lower than the WT uninjected embryos, it could be explained by the inevitable injection damage, which was also trivial. According to the data from National BioResource Project (https://www.shigen.nig.ac.jp/medaka/strain/hatchingRate.jsp), the inbred line HdrR-II show a hatching rate of 83%, which matches my results. Therefore, it can be concluded that the toxicity of MOMB to living medaka embryos is negligible.

4.3.3 Stability assay *in vivo*

The second parameter investigated *in vivo* was the stability of MOMB. The low toxicity of MOMB might be due to either a good biocompatibility or a low stability in developing embryos. As introduced in previous sections, the intracellular environment

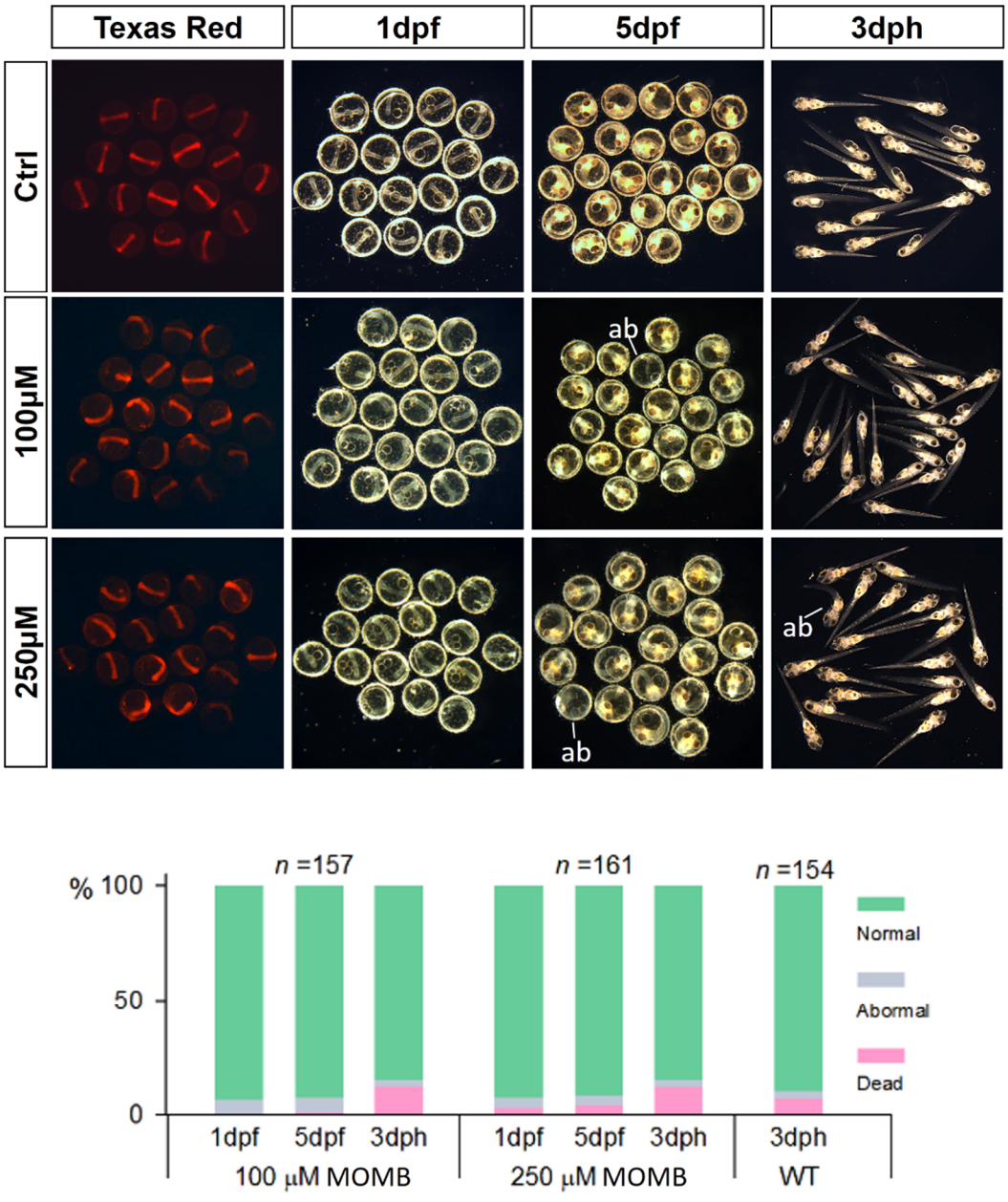


Figure 4-10. Toxicity of MOMB to embryonic development. Mixtures containing 5 mg/mL Texas Red, 5 mg/mL Texas Red with 100 μM or 250 μM MOMB were microinjected to 1-cell stage embryos. The toxicity of MOMB was assessed by counting dead and abnormal embryos over time of each injection group. Pictures shown are random samples from the injected embryos. Texas red channel indicates the homogeneity of the microinjection volume. ab, abnormal.

of developing embryos experience dramatic degradation of maternal mRNA in the early stages. Backbones that show nucleases resistance in solutions or cultured cells might fail the stability test in the stringent *in vivo* environment.

The long-term *in vivo* stability of MOMB is the prerequisite for its application in developing embryos. To determine the *in vivo* stability of MOMB, two approaches were used. The first approach is based on the fact that the degraded MOMB would lose its stem-loop conformation and release fluorophore, which exhibits as increased fluorescent signal. The second approach depends on LC-MS which identifies MOMB molecules from cell lysis samples.

4.3.3.1 Stability assay via fluorescence intensity

Thanks to the transparent and non-fluorescent nature of medaka embryos, the stability of MBs *in vivo* can be assessed by monitoring the fluorescent intensity changes of the injected MBs over time. To this end, mixtures of 5 mg/mL Texas Red dextran with either 100 μM MOMB or 100 μM DNAMB were injected into 1-cell stage embryos and the fluorescent intensity was monitored for the following 2 h. Data after 2 h were not used because as the embryos develop, the volume of cytoplasm increases rapidly and the fluorescence signal become diluted and vague. In addition, embryos injected with DNAMB showed severe deformation after 2 h, making it impossible for the observation. The degradation rate of MOMB and DNAMB were evaluated using the ratio of the FAM intensity to Texas Red intensity, termed G/R ratio for short. The fluorescent intensity from Texas Red served as the reference since the volume of the cytoplasm increases as the embryo develops. To guarantee the unbiased measurement of the fluorescent intensity, fluorescent images were taken under the same settings, and a representative square area with uniform fluorescent signal from the cytoplasm of each embryo was chosen to calculate the averaged intensity of green and red color

respectively via a python script written by myself (Appendix Python script 1). Images taken for the analysis are shown in Fig. 4-11 B, with dash frames indicating the chosen area. Although showing better performance *in vitro*, DNAMB had much higher background compared with MOMB right after injection without looking at the quantitative data, and the background increased gradually thereafter. Contrarily, MOMB showed a stable low background throughout 240 min. In the quantitative comparison as in Fig. 4-11 A, the G/R ratio of DNAMB increased steadily in the first 45 min and remain stable thereafter, indicating a quick degradation or disruption of DNAMB *in vivo*, which might be due to both the low ion strength intracellular environment that fail to keep the hydrogen bonds in the stem of DNAMB, and the various enzymes that compromise the stem-loop conformation of DNAMB. In contrast, the G/R ratio of MOMB remained the same in the 240 min assay, demonstrating the insensitivity of MOMB to both ion strength and intracellular enzymes. This preliminary assay proved the stability of MOMB *in vivo*.

4.3.3.2 Stability assays via LC-MS

To further investigate the long term stability of MOMB, LC-MS approach was used. First, a pure MOMB was analyzed via LC-MS to gain a reference profile. As in Fig. 4-12 A, 5 peaks were chosen as the signature profile of the MOMB. Then, MOMB was injected to 1-cell stage embryos. The injected embryos were allowed to develop for 6 h and 1 d followed by dechorionation and digestion with SDS and proteinase K for cell lysis preparation. The cell lysis from the injected embryos as well as from uninjected control embryos of 6 hpf were then subject to LC-MS analysis to investigate the existence of intact MOMB molecules. As in Fig. 4-12, the MS pattern of MOMB was absent in the 6 hpf control sample, while present in both injected samples, which proved the existence of MOMB for up to 1 day in living embryo. It

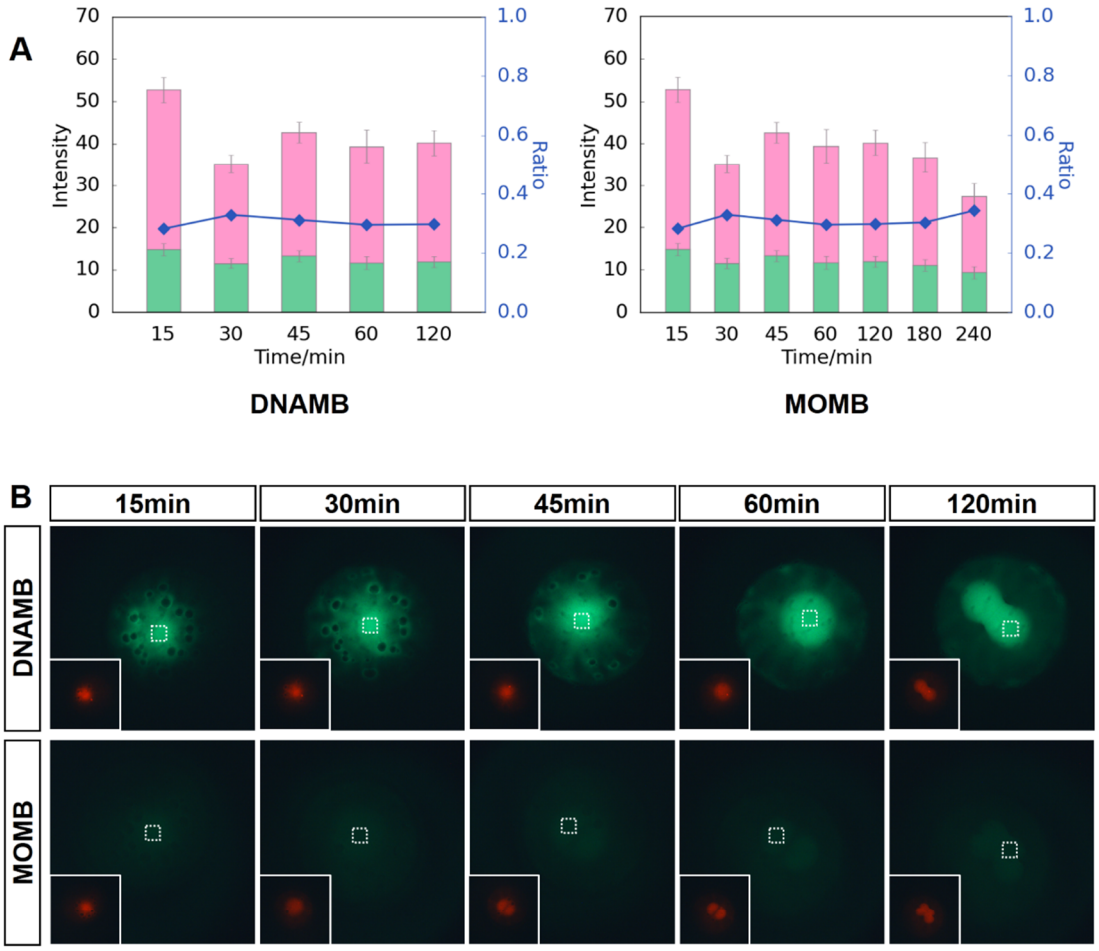


Figure 4-11. Stability of DNAMB and MOMB *in vivo* based on fluorescence change. Mixtures containing 5 mg/mL Texas Red and 50 μM DNAMB or 100 μM MOMB were microinjected to 1-cell stage embryos. Pictures were taken to assess the degradation rate over time of each MB. (A) FAM channel and Texas Red channel (lower left corner) of the injected embryos from 15 min to 120 min (DNAMB) or 240 min (MOMB) after injection. (B) Fluorescence intensity of each picture in (A) and the FAM/Texas Red ratio. The intensity was sampled pixel by pixel from the dash square frame in (A) and averaged.

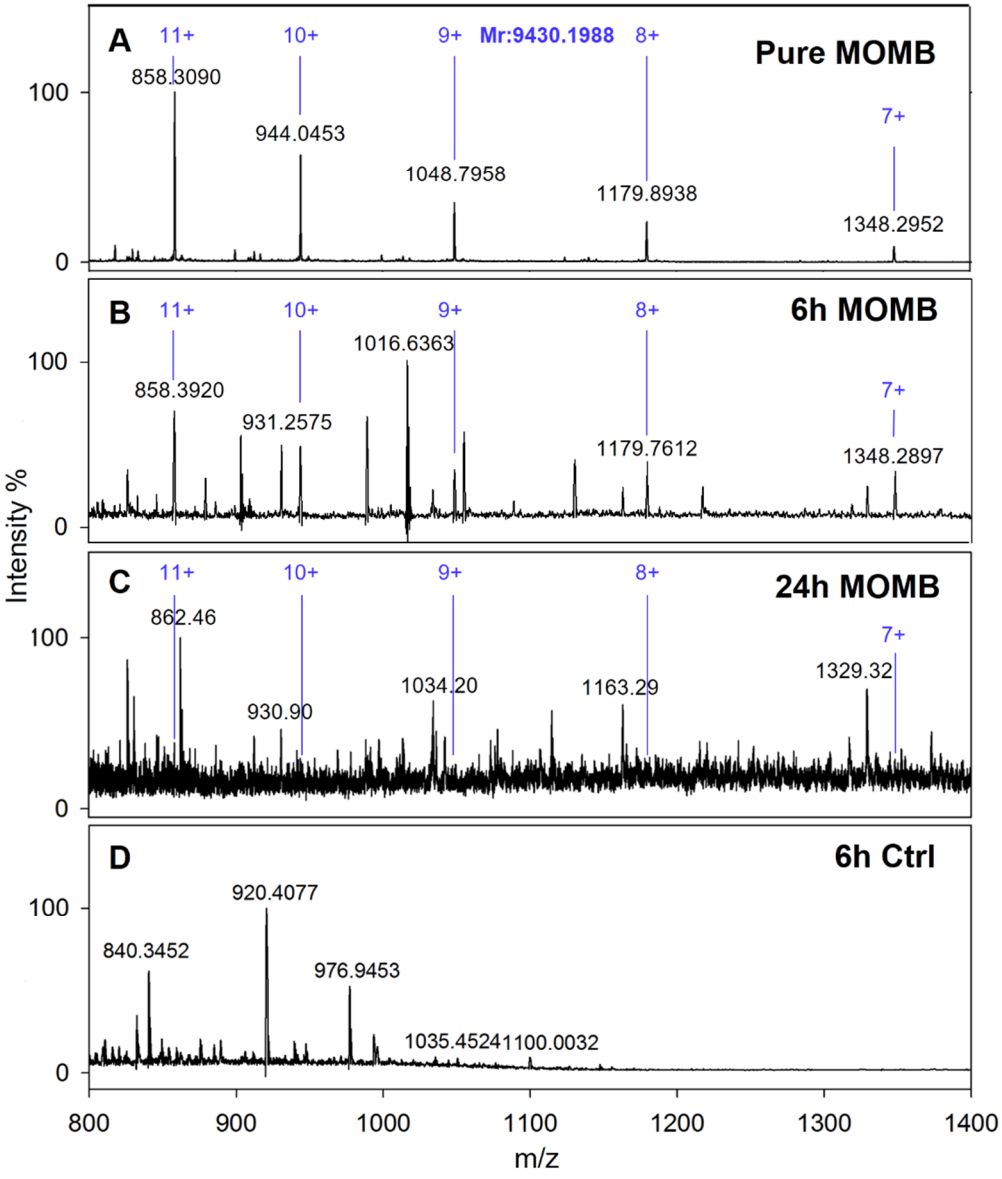


Figure 4-12. Stability of MOMB *in vivo* based on LC-MS. Mass spectrometry spectra of different samples are shown. (A) Pure MOMB sample showing the signature profile. (B) MOMB from embryos 6 h after injection. (C) MOMB from embryos 24 h after injection. (D) Cell lysis of control embryos.

suggested that MOMB has the potential to monitor RNA dynamics for at least one day. This result answered the question of how long can morpholino oligos sustain in living embryos. Though *in vitro* experiments demonstrated high resistance to nucleases of MOMBs, MOMBs still experienced substantial degradation in living embryos. To achieve a long-term visualization of RNA dynamics, further optimization of the MOMB probes is required.

4.3.4 Specificity assay *in vivo*

Finally, for an efficient RNA probe, specificity is also crucial. Though showing high specificity in PB solution, MOMB might bind to non-specific targets in developing embryos where over 50% of the genes show expression peaks (Aanes, Collas, and Aleström 2014; Arbeitman et al. 2002). It is necessary to verify the specificity of MOMB *in vivo*. To this end, MOMB was injected along with Texas Red as indicator to 1-cell stage embryos followed by the single-cell injection of perfect match (PM) or 2 nt mismatch (2MM) targets after three hours. Images were taken 10 min after the injection of the target. As shown in Fig.4-13, the Texas Red channel demonstrated the successful single cell microinjection of the MOMB, while the FAM channel showed the specific recognition signal of MOMB in response to only the PM target but not the 2MM target. It demonstrated the quick and precise response of MOMB to its target even under the stringent condition *in vivo*. Though the S/N ratio was much lower than the *in vitro* experiment, this is the first successful RNA visualization *in vivo*. The decreased S/N ratio might be due to the increased background signal from MOMB, as described in the previous section. Also, it could be due to the decreased hybridization rate of MOMB and target. It is possible that the short RNA targets were captured by certain binding proteins, which impedes their hybridization with MOMB.

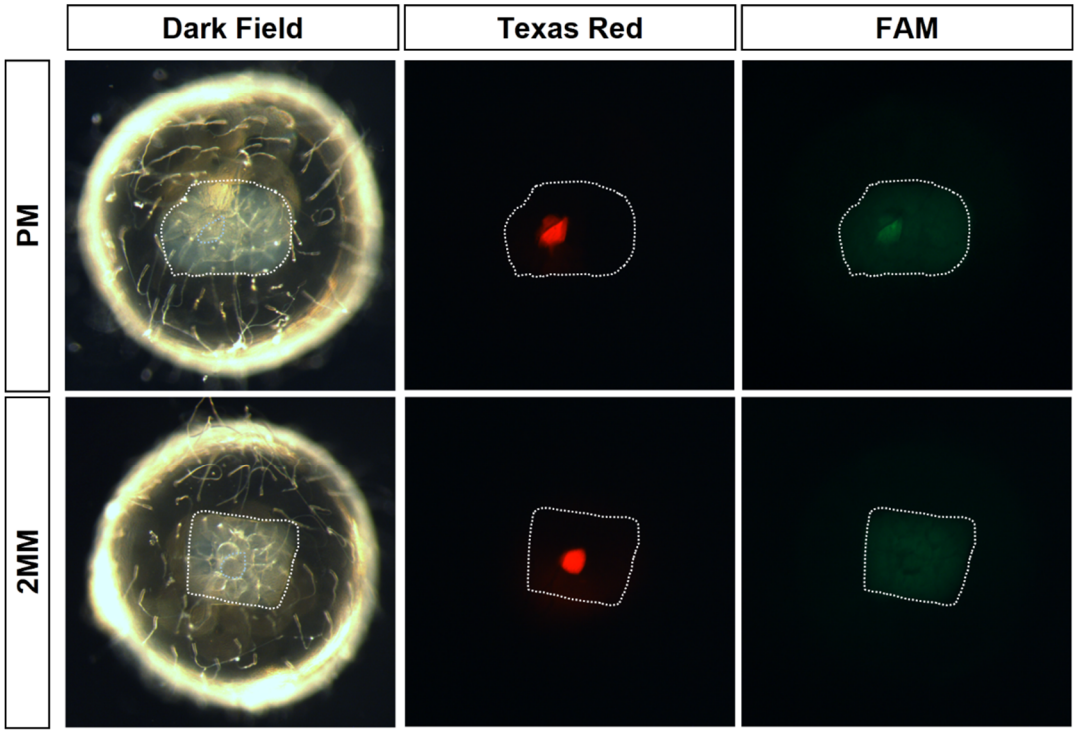


Figure 4-13. Specificity of MOMB *in vivo*. 200 μM MOMB was injected to 1-cell stage embryos. Three hours after injection, 10 μM perfect match (PM) target or 2-bp mismatch (2MM) target were mixed with 5mg/mL Texas Red and injected into one single cell.

CHAPTER 5 Conclusions and Perspectives

In this thesis, I proposed and established two new approaches to facilitate research in developmental biology. The first one aimed at rapid detection of subtle DNA alterations generated from gene editing experiments, while the second one aimed at non-invasive RNA visualization *in vivo*. Both approaches have proved to successfully reach the goal. Conclusions and perspectives of each approach is expounded below.

5.1 Detection of subtle DNA alterations by PAGE-HMA

In the first project of my thesis, I developed a rapid and cost-effective procedure for the detection, quantification and enrichment of the subtle DNA alterations produced by GE experiments. I showed that my PAGE-HMA procedure is a better method over other existing methods because it combined five major advantages. First, this method is straightforward without the requirement for enzymatic treatments or specific instruments, which facilitates its application in the normal biology labs. Second, PAGE-HMA has a high sensitivity by forming distinct Ht bands away from the bulk WT Hm bands, which allows the detection of rare alleles as low as 0.4%. Third, PAGE-HMA enables the identification and quantification of multiple alleles within a mixture. Four, the rare MT alleles can be enriched by gel recovery followed by one more round of PCR, which significantly reduces the workload for obtaining the sequence of the MT allele after cloning. Five, PAGE-HMA method also provides a tool to survey the GE target site for natural polymorphisms which might affect the result.

The PAGE-HMA method has certain limitations. Since PAGE-HMA only aims at rapid screening of the MT from GE experiments, it could fail to detect certain single base alterations. Nonetheless, for the majority of MT alleles, the alterations are usually more than one bp with certain deletions or insertions (Kim, Kweon, and Kim 2013), which can be clearly detected by PAGE-HMA. In addition, PAGE-HMA does not allow for high throughput screening. Detecting thousands of samples using PAGE-HMA approach would be laborious.

The PAGE-HMA method has been verified and used by other groups for gene targeting experiments (Ansai et al. 2014; Ota et al. 2013). Though developed for analyzing experimentally induced allelic alterations, the versatile PAGE-HMA method would find a wider variety of application in other areas for polymorphism detection in basic and applied research. For example, it has been utilized to facilitate species identification with our manuscript cited (Jung et al. 2014).

5.2 RNA visualization by MOMB *in vivo*

In the second project of my thesis, I proposed and constructed a new type of molecular beacons using the morpholino backbone, and demonstrated that MOMB has met the essential requirements for visualization of RNA targets *in vivo*. First, in the *in vitro* assays, the morpholino molecular beacon exhibited its insensitivity to salt concentration, high resistance to nucleases and DNA binding proteins, and high specificity in target recognition with the ability to discriminate sequence with 2 bp mismatch. I further applied this morpholino molecular beacon to living medaka embryos, where it showed negligible toxicity throughout the development process, stability for at least 1 day, and specificity to RNA targets *in vivo*. The above features make morpholino molecular beacon a prime candidate of the probes for RNA visualization *in vivo*.

MOMB targeting germ cell specific gene will be used to trace the migration of PGCs in preliminary experiments, but the background signal of MOMB might be too high for unambiguous observation. Further optimization could be made to overcome this limitations. For example, signal amplification strategies can be used to increase the S/N ratio. Such strategies could include multiplying the fluorophore and quencher of a single probe molecule, or adopting hairpin cascade reactions (C. Wu et al. 2015). To achieve long-term stability *in vivo*, hybrid backbones may be used for the molecular beacon with the incorporation of the more stable LNA or other nucleic acid analogs, so as to achieve a balance among the specificity, toxicity and stability of the MB probe.

REFERENCES

Aanes, Håvard, Philippe Collas, and Peter Aleström. 2014. “Transcriptome Dynamics and Diversity in the Early Zebrafish Embryo.” *Briefings in Functional Genomics* 13 (2): 95–105. doi:10.1093/bfgp/elt049.

Alvarez-Garcia, Ines, and Eric A. Miska. 2005. “MicroRNA Functions in Animal Development and Human Disease.” *Development* 132 (21): 4653–62. doi:10.1242/dev.02073.

Ansai, Satoshi, Keiji Inohaya, Yasutoshi Yoshiura, Manfred Schartl, Norihito Uemura, Ryosuke Takahashi, and Masato Kinoshita. 2014. “Design, Evaluation, and Screening Methods for Efficient Targeted Mutagenesis with Transcription Activator-like Effector Nucleases in Medaka.” *Development, Growth & Differentiation* 56 (1): 98–107. doi:10.1111/dgd.12104.

Arbeitman, Michelle N., Eileen E. M. Furlong, Farhad Imam, Eric Johnson, Brian H. Null, Bruce S. Baker, Mark A. Krasnow, Matthew P. Scott, Ronald W. Davis, and Kevin P. White. 2002. “Gene Expression During the Life Cycle of *Drosophila Melanogaster*.” *Science* 297 (5590): 2270–75. doi:10.1126/science.1072152.

Babendure, Jeremy R., Stephen R. Adams, and Roger Y. Tsien. 2003. “Aptamers Switch on Fluorescence of Triphenylmethane Dyes.” *Journal of the American Chemical Society* 125 (48): 14716–17. doi:10.1021/ja037994o.

Babon, J J, R Youil, and R G Cotton. 1995. “Improved Strategy for Mutation Detection--a Modification to the Enzyme Mismatch Cleavage Method.” *Nucleic Acids Research* 23 (24): 5082–84.

Bao, Gang, Won Jong Rhee, and Andrew Tsourkas. 2009. “Fluorescent Probes for Live-Cell RNA Detection.” *Annual Review of Biomedical Engineering* 11: 25–47. doi:10.1146/annurev-bioeng-061008-124920.

Bedell, Victoria M., Stephanie E. Westcot, and Stephen C. Ekker. 2011. “Lessons from Morpholino-Based Screening in Zebrafish.” *Briefings in Functional Genomics* 10 (4): 181–88. doi:10.1093/bfgp/elr021.

Beerli, Roger R., and Carlos F. Barbas. 2002. “Engineering Polydactyl Zinc-Finger Transcription Factors.” *Nature Biotechnology* 20 (2): 135–41. doi:10.1038/nbt0202-135.

Bertrand, Edouard, Pascal Chartrand, Matthias Schaefer, Shailesh M. Shenoy, Robert H. Singer, and Roy M. Long. 1998. “Localization of ASH1 mRNA Particles in Living Yeast.” *Molecular Cell* 2 (4): 437–45. doi:10.1016/S1097-2765(00)80143-4.

Bibikova, Marina, Dana Carroll, David J. Segal, Jonathan K. Trautman, Jeff Smith, Yang-Gyun Kim, and Srinivasan Chandrasegaran. 2001. “Stimulation of Homologous Recombination through Targeted Cleavage by Chimeric Nucleases.” *Molecular and Cellular Biology* 21 (1): 289–97. doi:10.1128/MCB.21.1.289-297.2001.

Birnboim, H. C., and J. Doly. 1979. “A Rapid Alkaline Extraction Procedure for Screening Recombinant Plasmid DNA.” *Nucleic Acids Research* 7 (6): 1513–23.

Boch, Jens, Heidi Scholze, Sebastian Schornack, Angelika Landgraf, Simone Hahn, Sabine Kay, Thomas Lahaye, Anja Nickstadt, and Ulla Bonas. 2009. “Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors.” *Science* 326 (5959): 1509–12. doi:10.1126/science.1178811.

Bonnet, Grégoire, Sanjay Tyagi, Albert Libchaber, and Fred Russell Kramer. 1999. “Thermodynamic Basis of the Enhanced Specificity of Structured DNA Probes.” *Proceedings of the National Academy of Sciences* 96 (11): 6171–76. doi:10.1073/pnas.96.11.6171.

Boutorine, Alexandre, Darya Novopashina, Olga Krasheninina, Karine Nozeret, and Alya Venyaminova. 2013. “Fluorescent Probes for Nucleic Acid Visualization in Fixed and Live Cells.” *Molecules* 18 (12): 15357–97. doi:10.3390/molecules181215357.

Bratu, Diana P., Byeong-Jik Cha, Musa M. Mhlanga, Fred Russell Kramer, and Sanjay Tyagi. 2003. “Visualizing the Distribution and Transport of mRNAs in Living Cells.” *Proceedings of the National Academy of Sciences of the United States of America* 100 (23): 13308–13. doi:10.1073/pnas.2233244100.

Buongiorno-Nardelli, M., and F. Amaldi. 1970. “Autoradiographic Detection of Molecular Hybrids between rRNA and DNA in Tissue Sections.” *Nature* 225 (5236): 946–48. doi:10.1038/225946a0.

Cabantous, Stéphanie, Hau B. Nguyen, Jean-Denis Pedelacq, Faten Koraïchi, Anu Chaudhary, Kumkum Ganguly, Meghan A. Lockard, Gilles Favre, Thomas C. Terwilliger, and Geoffrey S. Waldo. 2013. “A New Protein-Protein Interaction Sensor Based on Tripartite Split-GFP Association.” *Scientific Reports* 3 (October). doi:10.1038/srep02854.

Camp, Esther, Ana V. Sánchez-Sánchez, Antonio García-España, Rob DeSalle, Lina Odqvist, José Enrique O’Connor, and José L. Mullor. 2009. “*Nanog* Regulates Proliferation During Early Fish Development.” *STEM CELLS* 27 (9): 2081–91. doi:10.1002/stem.133.

Carlson, Bruce M. 2013. *Human Embryology and Developmental Biology*. Elsevier Health Sciences.

Carroll, Dana. 2011. “Genome Engineering With Zinc-Finger Nucleases.” *Genetics* 188 (4): 773–82. doi:10.1534/genetics.111.131433.

Catrina, Irina E., Salvatore A. E. Marras, and Diana P. Bratu. 2012. “Tiny Molecular Beacons: LNA/2′-O-Methyl RNA Chimeric Probes for Imaging Dynamic mRNA Processes in Living Cells.” *ACS Chemical Biology* 7 (9): 1586–95. doi:10.1021/cb300178a.

Chakravarti, D., P. Mailander, J. Franzen, S. Higginbotham, E. L. Cavalieri, and E. G. Rogan. 1998. “Detection of Dibenzo[a,l]pyrene-Induced H-Ras Codon 61 Mutant Genes in Preneoplastic SENCAR Mouse Skin Using a New PCR-RFLP Method.” *Oncogene* 16 (24): 3203–10. doi:10.1038/sj.onc.1201853.

Cheng, Jinping, Yongyou Zhang, and Qingge Li. 2004. “Real-Time PCR Genotyping Using Displacing Probes.” *Nucleic Acids Research* 32 (7): e61. doi:10.1093/nar/gnh055.

Chen, Jianbin, Xi Zhang, Tiansu Wang, Zhendong Li, Guijun Guan, and Yunhan Hong. 2012. “Efficient Detection, Quantification and

Enrichment of Subtle Allelic Alterations." *DNA Research* 19 (5): 423–33. doi:10.1093/dnares/dss023.

Coonrod, Scott A., Laura C. Bolling, Paul W. Wright, Pablo E. Visconti, and John C. Herr. 2001. "A Morpholino Phenocopy of the Mouse MOS Mutation." *Genesis* 30 (3): 198–200. doi:10.1002/gene.1065.

Dahlem, Timothy J., Kazuyuki Hoshijima, Michael J. Jurynec, Derrick Gunther, Colby G. Starker, Alexandra S. Locke, Allison M. Weis, Daniel F. Voytas, and David Jonah Grunwald. 2012. "Simple Methods for Generating and Detecting Locus-Specific Mutations Induced with TALENs in the Zebrafish Genome." *PLoS Genetics* 8 (8). doi:10.1371/journal.pgen.1002861.

Daigle, Nathalie, and Jan Ellenberg. 2007. "λN-GFP: An RNA Reporter System for Live-Cell Imaging." *Nature Methods* 4 (8): 633–36. doi:10.1038/nmeth1065.

Dean, Kevin M., and Amy E. Palmer. 2014. "Advances in Fluorescence Labeling Strategies for Dynamic Cellular Imaging." *Nature Chemical Biology* 10 (7): 512–23. doi:10.1038/nchembio.1556.

Deriano, Ludovic, and David B. Roth. 2013. "Modernizing the Nonhomologous End-Joining Repertoire: Alternative and Classical NHEJ Share the Stage." *Annual Review of Genetics* 47: 433–55. doi:10.1146/annurev-genet-110711-155540.

Dolgosheina, Elena V., Sunny C. Y. Jeng, Shanker Shyam S. Panchapakesan, Razvan Cojocaru, Patrick S. K. Chen, Peter D. Wilson, Nancy Hawkins, Paul A. Wiggins, and Peter J. Unrau. 2014. "RNA Mango Aptamer-Fluorophore: A Bright, High-Affinity Complex for RNA Labeling and Tracking." *ACS Chemical Biology* 9 (10): 2412–20. doi:10.1021/cb500499x.

Doyle, E. L., N. J. Booher, D. S. Standage, D. F. Voytas, V. P. Brendel, J. K. VanDyk, and A. J. Bogdanove. 2012. "TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: Tools for TAL Effector Design and Target Prediction." *Nucleic Acids Research* 40 (W1): W117–22. doi:10.1093/nar/gks608.

Ekker, Stephen C., and Jon D. Larson. 2001. "Morphant Technology in Model Developmental Systems." *Genesis* 30 (3): 89–93. doi:10.1002/gene.1038.

Feng, Zhengyan, Botao Zhang, Wona Ding, Xiaodong Liu, Dong-Lei Yang, Pengliang Wei, Fengqiu Cao, et al. 2013. "Efficient Genome Editing in Plants Using a CRISPR/Cas System." *Cell Research* 23 (10): 1229–32. doi:10.1038/cr.2013.114.

Ferg, Marco, Remo Sanges, Jochen Gehrig, Janos Kiss, Matthias Bauer, Agnes Lovas, Monika Szabo, et al. 2007. "The TATA-Binding Protein Regulates Maternal mRNA Degradation and Differential Zygotic Transcription in Zebrafish." *The EMBO Journal* 26 (17): 3945–56. doi:10.1038/sj.emboj.7601821.

Fineran, Peter C., and Emmanuelle Charpentier. 2012. "Memory of Viral Infections by CRISPR-Cas Adaptive Immune Systems: Acquisition of New Information." *Virology* 434 (2): 202–9. doi:10.1016/j.virol.2012.10.003.

Foley, Jonathan E., Morgan L. Maeder, Joseph Pearlberg, J. Keith Joung, Randall T. Peterson, and Jing-Ruey J. Yeh. 2009. "Targeted

Mutagenesis in Zebrafish Using Customized Zinc-Finger Nucleases." *Nature Protocols* 4 (12): 1855–67. doi:10.1038/nprot.2009.209.

Furutani-Seiki, Makoto, Takao Sasado, Chikako Morinaga, Hiroshi Suwa, Katsutoshi Niwa, Hiroki Yoda, Tomonori Deguchi, et al. 2004. "A Systematic Genome-Wide Screen for Mutations Affecting Organogenesis in Medaka, *Oryzias latipes*." *Mechanisms of Development*, Medaka, 121 (7–8): 647–58. doi:10.1016/j.mod.2004.04.016.

Fusco, Dahlene, Nathalie Accornero, Brigitte Lavoie, Shailesh M. Shenoy, Jean Marie Blanchard, Robert H. Singer, and Edouard Bertrand. 2003. "Single mRNA Molecules Demonstrate Probabilistic Movement in Living Mammalian Cells." *Current Biology: CB* 13 (2): 161–67.

Gaj, Thomas, Charles A. Gersbach, and Carlos F. Barbas. 2013. "ZFN, TALEN, and CRISPR/Cas-Based Methods for Genome Engineering." *Trends in Biotechnology* 31 (7): 397–405. doi:10.1016/j.tibtech.2013.04.004.

Gall, Joseph G., and Mary Lou Pardue. 1969. "Formation and Detection of RNA-DNA Hybrid Molecules in Cytological Preparations." *Proceedings of the National Academy of Sciences* 63 (2): 378–83.

Gavis, Elizabeth R., and Ruth Lehmann. 1992. "Localization of Nanos RNA Controls Embryonic Polarity." *Cell* 71 (2): 301–13. doi:10.1016/0092-8674(92)90358-J.

Gerety, Sebastian S., and David G. Wilkinson. 2011. "Morpholino Artifacts Provide Pitfalls and Reveal a Novel Role for pro-Apoptotic Genes in Hindbrain Boundary Development." *Developmental Biology* 350 (2): 279–89. doi:10.1016/j.ydbio.2010.11.030.

Glavac, D., and M. Dean. 1995. "Applications of Heteroduplex Analysis for Mutation Detection in Disease Genes." *Human Mutation* 6 (4): 281–87. doi:10.1002/humu.1380060402.

Grünwald, David, and Robert H. Singer. 2010. "In Vivo Imaging of Labelled Endogenous B-Actin mRNA during Nucleocytoplasmic Transport." *Nature* 467 (7315): 604–7. doi:10.1038/nature09438.

Hamatani, Toshio, Mark G. Carter, Alexei A. Sharov, and Minoru S. H. Ko. 2004. "Dynamics of Global Gene Expression Changes during Mouse Preimplantation Development." *Developmental Cell* 6 (1): 117–31.

Heasman, Janet, Matt Kofron, and Chris Wylie. 2000. "βCatenin Signaling Activity Dissected in the Early *Xenopus* Embryo: A Novel Antisense Approach." *Developmental Biology* 222 (1): 124–34. doi:10.1006/dbio.2000.9720.

Hermann, Thomas, and Dinshaw J. Patel. 2000. "Adaptive Recognition by Nucleic Acid Aptamers." *Science* 287 (5454): 820–25. doi:10.1126/science.287.5454.820.

Highsmith, W. Edward, Qian Jin, Arun J. Nataraj, Jacquelyn M. O'Connor, Valerie D. Burland, Wendy R. Baubonis, Foner P. Curtis, Noriko Kusukawa, and Mark M. Garner. 1999. "Use of a DNA Toolbox for the Characterization of Mutation Scanning Methods. I: Construction of the Toolbox and Evaluation of Heteroduplex Analysis." *Electrophoresis* 20 (6): 1186–94. doi:10.1002/(SICI)1522-2683(19990101)20:6<1186::AID-ELPS1186>3.0.CO;2-6.

Hong, Yunhan, Tongming Liu, Haobin Zhao, Hongyan Xu, Weijia Wang, Rong Liu, Tiansheng Chen, Jiaorong Deng, and Jianfang Gui. 2004. “Establishment of a Normal Medakafish Spermatogonial Cell Line Capable of Sperm Production *in Vitro*.” *Proceedings of the National Academy of Sciences of the United States of America* 101 (21): 8011–16. doi:10.1073/pnas.0308668101.

Hong, Yunhan, Christoph Winkler, and Manfred Schartl. 1998. “Production of Medakafish Chimeras from a Stable Embryonic Stem Cell Line.” *Proceedings of the National Academy of Sciences* 95 (7): 3679–84.

Hong, Y., C. Winkler, and M. Schartl. 1996. “Pluripotency and Differentiation of Embryonic Stem Cell Lines from the Medakafish (*Oryzias latipes*).” *Mechanisms of Development* 60 (1): 33–44.

Hoozemans, Diederik A, Roel Schats, Cornelis B Lambalk, Roy Homburg, and Peter GA Hompes. 2004. “Human Embryo Implantation: Current Knowledge and Clinical Implications in Assisted Reproductive Technology.” *Reproductive BioMedicine Online* 9 (6): 692–715. doi:10.1016/S1472-6483(10)61781-6.

Howard, E. W., L. A. Newman, D. W. Oleksyn, R. C. Angerer, and L. M. Angerer. 2001. “SpKrl: A Direct Target of Beta-Catenin Regulation Required for Endoderm Differentiation in Sea Urchin Embryos.” *Development* 128 (3): 365–75.

Hruscha, Alexander, Peter Krawitz, Alexandra Rechenberg, Verena Heinrich, Jochen Hecht, Christian Haass, and Bettina Schmid. 2013. “Efficient CRISPR/Cas9 Genome Editing with Low off-Target Effects in Zebrafish.” *Development* 140 (24): 4982–87. doi:10.1242/dev.099085.

Hudziak, R. M., E. Barofsky, D. F. Barofsky, D. L. Weller, S. B. Huang, and D. D. Weller. 1996. “Resistance of Morpholino Phosphorodiamidate Oligomers to Enzymatic Degradation.” *Antisense & Nucleic Acid Drug Development* 6 (4): 267–72.

Hyrup, Birgitte, and Peter E. Nielsen. 1996. “Peptide Nucleic Acids (PNA): Synthesis, Properties and Potential Applications.” *Bioorganic & Medicinal Chemistry* 4 (1): 5–23. doi:10.1016/0968-0896(95)00171-9.

Ingber, Donald E., and Michael Levin. 2007. “What Lies at the Interface of Regenerative Medicine and Developmental Biology?” *Development* 134 (14): 2541–47. doi:10.1242/dev.003707.

Itzkovitz, Shalev, and Alexander van Oudenaarden. 2011. “Validating Transcripts with Probes and Imaging Technology.” *Nature Methods* 8 (4 Suppl): S12–19. doi:10.1038/nmeth.1573.

Iwamatsu, Takashi. 2004. “Stages of Normal Development in the Medaka *Oryzias latipes*.” *Mechanisms of Development* 121 (7–8): 605–18. doi:10.1016/j.mod.2004.03.012.

Jao, Li-En, Susan R. Wente, and Wenbiao Chen. 2013. “Efficient Multiplex Biallelic Zebrafish Genome Editing Using a CRISPR Nuclease System.” *Proceedings of the National Academy of Sciences* 110 (34): 13904–9. doi:10.1073/pnas.1308335110.

Jasin, Maria. 1996. “Genetic Manipulation of Genomes with Rare-Cutting Endonucleases.” *Trends in Genetics* 12 (6): 224–28. doi:10.1016/0168-9525(96)10019-6.

John, H. A., M. L. Birnstiel, and K. W. Jones. 1969. “RNA-DNA Hybrids at the Cytological Level.” *Nature* 223 (5206): 582–87. doi:10.1038/223582a0.

Joung, J. Keith, and Jeffry D. Sander. 2012. “TALENs: A Widely Applicable Technology for Targeted Genome Editing.” *Nature Reviews Molecular Cell Biology* 14 (November): 49–55. doi:10.1038/nrm3486.

Jung, Juyeon, Kyung Hee Kim, Kiwoung Yang, Kyong-Hwan Bang, and Tae-Jin Yang. 2014. “Practical Application of DNA Markers for High-Throughput Authentication of *Panax Ginseng* and *Panax Quinquefolius* from Commercial Ginseng Products.” *Journal of Ginseng Research* 38 (2): 123–29. doi:10.1016/j.jgr.2013.11.017.

Kam, Yossi, Abraham Rubinstein, Aviram Nissan, David Halle, and Eylon Yavin. 2012. “Detection of Endogenous K-Ras mRNA in Living Cells at a Single Base Resolution by a PNA Molecular Beacon.” *Molecular Pharmaceutics* 9 (3): 685–93. doi:10.1021/mp200505k.

Karkare, Shantanu, and Deepak Bhatnagar. 2006. “Promising Nucleic Acid Analogs and Mimics: Characteristic Features and Applications of PNA, LNA, and Morpholino.” *Applied Microbiology and Biotechnology* 71 (5): 575–86. doi:10.1007/s00253-006-0434-2.

Kasahara, Masahiro, Kiyoshi Naruse, Shin Sasaki, Yoichiro Nakatani, Wei Qu, Budrul Ahsan, Tomoyuki Yamada, et al. 2007. “The Medaka Draft Genome and Insights into Vertebrate Genome Evolution.” *Nature* 447 (7145): 714–19. doi:10.1038/nature05846.

Kim, Yongsub, Jiyeon Kweon, and Jin-Soo Kim. 2013. “TALENs and ZFNs Are Associated with Different Mutation Signatures.” *Nature Methods* 10 (3): 185–185. doi:10.1038/nmeth.2364.

Kinoshita, Masato, Kenji Murata, Kiyoshi Naruse, and Minoru Tanaka. 2009. “Medaka Management.” In *Medaka*, 31–66. John Wiley & Sons, Ltd. http://onlinelibrary.wiley.com.libproxy1.nus.edu.sg/doi/10.1002/9780813818849.ch2/summary.

Kirchmaier, Stephan, Burkhard Höckendorf, Eva Katharina Möller, Dorothee Bornhorst, Francois Spitz, and Joachim Wittbrodt. 2013. “Efficient Site-Specific Transgenesis and Enhancer Activity Tests in Medaka Using PhiC31 Integrase.” *Development* 140 (20): 4287–95. doi:10.1242/dev.096081.

Kirchmaier, Stephan, Kiyoshi Naruse, Joachim Wittbrodt, and Felix Loosli. 2015. “The Genomic and Genetic Toolbox of the Teleost Medaka (*Oryzias latipes*).” *Genetics* 199 (4): 905–18. doi:10.1534/genetics.114.173849.

Kloosterman, Wigard P., Erno Wienholds, Ewart de Bruijn, Sakari Kauppinen, and Ronald H. A. Plasterk. 2006. “In Situ Detection of miRNAs in Animal Embryos Using LNA-Modified Oligonucleotide Probes.” *Nature Methods* 3 (1): 27–29. doi:10.1038/nmeth843.

Kos, R., M. V. Reedy, R. L. Johnson, and C. A. Erickson. 2001. “The Winged-Helix Transcription Factor FoxD3 Is Important for Establishing the Neural Crest Lineage and Repressing Melanogenesis in Avian Embryos.” *Development* 128 (8): 1467–79.

Kostrikis, L. G., E. Bagdades, Y. Cao, L. Zhang, D. Dimitriou, and D. D. Ho. 1995. “Genetic Analysis of Human Immunodeficiency Virus Type 1 Strains from Patients in Cyprus: Identification of a New Subtype Designated Subtype I.” *Journal of Virology* 69 (10): 6122.

Kuhn, Heiko, Vadim V. Demidov, James M. Coull, Mark J. Fiandaca, Brian D. Gildea, and Maxim D. Frank-Kamenetskii. 2002. “Hybridization of DNA

and PNA Molecular Beacons to Single-Stranded and Double-Stranded DNA Targets." *Journal of the American Chemical Society* 124 (6): 1097–1103. doi:10.1021/ja0041324.

Larson, Daniel R., Daniel Zenklusen, Bin Wu, Jeffrey A. Chao, and Robert H. Singer. 2011. "Real-Time Observation of Transcription Initiation and Elongation on an Endogenous Yeast Gene." *Science* 332 (6028): 475–78. doi:10.1126/science.1202142.

Leys, Eugene J., James H. Smith, Sharon R. Lyons, and Ann L. Griffen. 1999. "Identification of *Porphyromonas Gingivalis* Strains by Heteroduplex Analysis and Detection of Multiple Strains." *Journal of Clinical Microbiology* 37 (12): 3906–11.

Li, Jianwei Jeffery, Xiaohong Fang, Sheldon M. Schuster, and Weihong Tan. 2000. "Molecular Beacons: A Novel Approach to Detect Protein – DNA Interactions." *Angewandte Chemie International Edition* 39 (6): 1049–52. doi:10.1002/(SICI)1521-3773(20000317)39:6<1049::AID-ANIE1049>3.0.CO;2-2.

Li, Jianwei Jeffery, Ron Geyer, and Weihong Tan. 2000. "Using Molecular Beacons as a Sensitive Fluorescence Assay for Enzymatic Cleavage of Single-Stranded DNA." *Nucleic Acids Research* 28 (11): e52.

Li, Mingyou, Ni Hong, Hongyan Xu, Meisheng Yi, Changming Li, Jianfang Gui, and Yunhan Hong. 2009. "Medaka *Vasa* Is Required for Migration but Not Survival of Primordial Germ Cells." *Mechanisms of Development* 126 (5–6): 366–81. doi:10.1016/j.mod.2009.02.004.

Lionnet, Timothée, Kevin Czaplinski, Xavier Darzacq, Yaron Shav-Tal, Amber L. Wells, Jeffrey A. Chao, Hye Yoon Park, Valeria de Turris, Melissa Lopez-Jones, and Robert H. Singer. 2011. "A Transgenic Mouse for in Vivo Detection of Endogenous Labeled mRNA." *Nature Methods* 8 (2): 165–70. doi:10.1038/nmeth.1551.

Li, Qingge, Guoyan Luan, Qiuping Guo, and Jixuan Liang. 2002. "A New Class of Homogeneous Nucleic Acid Probes Based on Specific Displacement Hybridization." *Nucleic Acids Research* 30 (2): e5.

Majlessi, M., N. C. Nelson, and M. M. Becker. 1998. "Advantages of 2'-O-Methyl Oligoribonucleotide Probes for Detecting RNA Targets." *Nucleic Acids Research* 26 (9): 2224–29.

Markham, Nicholas R., and Michael Zuker. 2005. "DINAMelt Web Server for Nucleic Acid Melting Prediction." *Nucleic Acids Research* 33 (Web Server issue): W577–81. doi:10.1093/nar/gki591.

———. 2008. "UNAFold: Software for Nucleic Acid Folding and Hybridization." *Methods in Molecular Biology (Clifton, N.J.)* 453: 3–31. doi:10.1007/978-1-60327-429-6_1.

Mathavan, Sinnakaruppan, Serene G. P Lee, Alicia Mak, Lance D Miller, Karuturi Radha Krishna Murthy, Kunde R Govindarajan, Yan Tong, et al. 2005. "Transcriptome Analysis of Zebrafish Embryogenesis Using Microarrays." *PLoS Genet* 1 (2): e29. doi:10.1371/journal.pgen.0010029.

Medley, Colin D., and Zhi Zhu. 2013. "Molecular Beacons for Intracellular Analysis: Challenges and Successes." In *Molecular Beacons*, edited by Chaoyong James Yang and Weihong Tan, 153–74. Springer Berlin Heidelberg.

http://link.springer.com.libproxy1.nus.edu.sg/chapter/10.1007/978-3-642-39109-5_10.

Meng, Xiangdong, Marcus B. Noyes, Lihua J. Zhu, Nathan D. Lawson, and Scot A. Wolfe. 2008. “Targeted Gene Inactivation in Zebrafish Using Engineered Zinc-Finger Nucleases.” *Nature Biotechnology* 26 (6): 695–701. doi:10.1038/nbt1398.

Miller, Jeffrey C., Siyuan Tan, Guijuan Qiao, Kyle A. Barlow, Jianbin Wang, Danny F. Xia, Xiangdong Meng, et al. 2011. “A TALE Nuclease Architecture for Efficient Genome Editing.” *Nature Biotechnology* 29 (2): 143–48. doi:10.1038/nbt.1755.

Molenaar, C., S. A. Marras, J. C. M. Slats, J.-C. Truffert, M. Lemaître, A. K. Raap, R. W. Dirks, and H. J. Tanke. 2001. “Linear 2′ O-Methyl RNA Probes for the Visualization of RNA in Living Cells.” *Nucleic Acids Research* 29 (17): e89–e89. doi:10.1093/nar/29.17.e89.

Monroy-Contreras, Ricardo, and Luis Vaca. 2011. “Molecular Beacons: Powerful Tools for Imaging RNA in Living Cells.” *Journal of Nucleic Acids* 2011 (August): e741723. doi:10.4061/2011/741723.

Moore, J K, and J E Haber. 1996. “Cell Cycle and Genetic Requirements of Two Pathways of Nonhomologous End-Joining Repair of Double-Strand Breaks in Saccharomyces Cerevisiae.” *Molecular and Cellular Biology* 16 (5): 2164–73.

Morvan, F., H. Porumb, G. Degols, I. Lefebvre, A. Pompon, B. S. Sproat, B. Rayner, C. Malvy, B. Lebleu, and J. L. Imbach. 1993. “Comparative Evaluation of Seven Oligonucleotide Analogs as Potential Antisense Agents.” *Journal of Medicinal Chemistry* 36 (2): 280–87. doi:10.1021/jm00054a013.

Moscou, Matthew J., and Adam J. Bogdanove. 2009. “A Simple Cipher Governs DNA Recognition by TAL Effectors.” *Science* 326 (5959): 1501–1501. doi:10.1126/science.1178817.

Nasevicius, Aidas, and Stephen C. Ekker. 2000. “Effective Targeted Gene ‘knockdown’ in Zebrafish.” *Nature Genetics* 26 (2): 216–20. doi:10.1038/79951.

Nataraj, Arun J., Isabelle Olivos-Glander, Noriko Kusukawa, and W. Edward Highsmith. 1999. “Single-Strand Conformation Polymorphism and Heteroduplex Analysis for Gel-Based Mutation Detection.” *ELECTROPHORESIS* 20 (6): 1177–85. doi:10.1002/(SICI)1522-2683(19990101)20:6<1177::AID-ELPS1177>3.0.CO;2-2.

Nielsen, P. E., and M. Egholm. 1999. “An Introduction to Peptide Nucleic Acid.” *Current Issues in Molecular Biology* 1 (1-2): 89–104.

Nutt, S. L., O. J. Bronchain, K. O. Hartley, and E. Amaya. 2001. “Comparison of Morpholino Based Translational Inhibition during the Development of *Xenopus Laevis* and *Xenopus Tropicalis*.” *Genesis (New York, N.Y.: 2000)* 30 (3): 110–13.

OECD. 2013. *Test No. 236: Fish Embryo Acute Toxicity (FET) Test*. Paris: Organisation for Economic Co-operation and Development. http://www.oecd-ilibrary.org/content/book/9789264203709-en.

Oleykowski, C A, C R Bronson Mullins, A K Godwin, and A T Yeung. 1998. “Mutation Detection Using a Novel Plant Endonuclease.” *Nucleic Acids Research* 26 (20): 4597–4602.

Ortiz, E, G Estrada, and P. M Lizardi. 1998. “PNA Molecular Beacons for Rapid Detection of PCR Amplicons.” *Molecular and Cellular Probes* 12 (4): 219–26. doi:10.1006/mcpr.1998.0175.

Ota, Satoshi, Yu Hisano, Michiko Muraki, Kazuyuki Hoshijima, Timothy J. Dahlem, David J. Grunwald, Yasushi Okada, and Atsuo Kawahara. 2013. “Efficient Identification of TALEN-Mediated Genome Modifications Using Heteroduplex Mobility Assays.” *Genes to Cells* 18 (6): 450–58. doi:10.1111/gtc.12050.

Ouyang, Xiaohu, Ilya A. Shestopalov, Surajit Sinha, Genhua Zheng, Cameron L. W. Pitt, Wen-Hong Li, Andrew J. Olson, and James K. Chen. 2009. “Versatile Synthesis and Rational Design of Caged Morpholinos.” *Journal of the American Chemical Society* 131 (37): 13255–69. doi:10.1021/ja809933h.

Paige, Jeremy S., Karen Y. Wu, and Samie R. Jaffrey. 2011. “RNA Mimics of Green Fluorescent Protein.” *Science* 333 (6042): 642–46. doi:10.1126/science.1207339.

Pantazis, Periklis, and Willy Supatto. 2014. “Advances in Whole-Embryo Imaging: A Quantitative Transition Is Underway.” *Nature Reviews Molecular Cell Biology* 15 (5): 327–39. doi:10.1038/nrm3786.

Parant, John M., Stephen A. George, Rob Pryor, Carl T. Wittwer, and H. Joseph Yost. 2009. “A Rapid and Efficient Method of Genotyping Zebrafish Mutants.” *Developmental Dynamics* 238 (12): 3168–74. doi:10.1002/dvdy.22143.

Park, Hye Yoon, Hyungsik Lim, Young J. Yoon, Antonia Follenzi, Chiso Nwokafor, Melissa Lopez-Jones, Xiuhua Meng, and Robert H. Singer. 2014. “Visualization of Dynamics of Single Endogenous mRNA Labeled in Live Mouse.” *Science* 343 (6169): 422–24. doi:10.1126/science.1239200.

Piatek, Amy S., Sanjay Tyagi, Arno C. Pol, Amalio Telenti, Lincoln P. Miller, Fred Russell Kramer, and David Alland. 1998. “Molecular Beacon Sequence Analysis for Detecting Drug Resistance in Mycobacterium Tuberculosis.” *Nature Biotechnology* 16 (4): 359–63. doi:10.1038/nbt0498-359.

Pogorelov, Alexander G., Igor I. Katkov, Evgenia I. Smolyaninova, and Dmitri V. Goldshtein. 2006. “Changes in Intracellular Potassium and Sodium Content of 2-Cell Mouse Embryos Induced by Exposition to Vitrification Concentrations of Ethylene Glycol.” *Cryoletters* 27 (2): 87–98.

Porazinski, Sean R., Huijia Wang, and Makoto Furutani-Seiki. 2010. “Microinjection of Medaka Embryos for Use as a Model Genetic Organism.” *Journal of Visualized Experiments*, no. 46 (December). doi:10.3791/1937.

———. 2011. “Essential Techniques for Introducing Medaka to a Zebrafish Laboratory--towards the Combined Use of Medaka and Zebrafish for Further Genetic Dissection of the Function of the Vertebrate Genome.” *Methods in Molecular Biology (Clifton, N.J.)* 770: 211–41. doi:10.1007/978-1-61779-210-6_8.

Porazinski, Sean, Huijia Wang, Yoichi Asaoka, Martin Behrndt, Tatsuo Miyamoto, Hitoshi Morita, Shoji Hata, et al. 2015. “YAP Is Essential for Tissue Tension to Ensure Vertebrate 3D Body Shape.” *Nature* 521 (7551): 217–21. doi:10.1038/nature14215.

Porteus, Matthew H., and David Baltimore. 2003. “Chimeric Nucleases Stimulate Gene Targeting in Human Cells.” *Science* 300 (5620): 763–763. doi:10.1126/science.1078395.

Pourzand, Charareh, and Peter Cerutti. 1993. “Genotypic Mutation Analysis by RFLP/PCR.” *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 288 (1): 113–21. doi:10.1016/0027-5107(93)90213-Y.

Qiu, Peter, Harini Shandilya, James M. D’Alessio, Kevin O’Connor, Jeffrey Durocher, and Gary F. Gerard. 2004. “Mutation Detection Using Surveyor Nuclease.” *BioTechniques* 36 (4): 702–7.

Rackham, Oliver, and Chris M Brown. 2004. “Visualization of RNA–protein Interactions in Living Cells: FMRP and IMP1 Interact on mRNAs.” *The EMBO Journal* 23 (16): 3346–55. doi:10.1038/sj.emboj.7600341.

Rebagliati, M. R., D. L. Weeks, R. P. Harvey, and D. A. Melton. 1985. “Identification and Cloning of Localized Maternal RNAs from *Xenopus* Eggs.” *Cell* 42 (3): 769–77. doi:10.1016/0092-8674(85)90273-9.

Rembold, Martina, Kajori Lahiri, Nicholas S. Foulkes, and Joachim Wittbrodt. 2006. “Transgenesis in Fish: Efficient Selection of Transgenic Fish by Co-Injection with a Fluorescent Reporter Construct.” *Nature Protocols* 1 (3): 1133–39. doi:10.1038/nprot.2006.165.

Reyon, Deepak, Shengdar Q. Tsai, Cyd Khayter, Jennifer A. Foden, Jeffry D. Sander, and J. Keith Joung. 2012. “FLASH Assembly of TALENs for High-Throughput Genome Editing.” *Nature Biotechnology* 30 (5): 460–65. doi:10.1038/nbt.2170.

Ross, Jeffrey J., Osamu Shimmi, Peter Vilmos, Anna Petryk, Hyon Kim, Karin Gaudenz, Spencer Hermanson, Stephen C. Ekker, Michael B. O’Connor, and J. Lawrence Marsh. 2001. “Twisted Gastrulation Is a Conserved Extracellular BMP Antagonist.” *Nature* 410 (6827): 479–83. doi:10.1038/35068578.

Rouet, P., F. Smih, and M. Jasin. 1994. “Introduction of Double-Strand Breaks into the Genome of Mouse Cells by Expression of a Rare-Cutting Endonuclease.” *Molecular and Cellular Biology* 14 (12): 8096–8106.

Ruano, G., and K. K. Kidd. 1992. “Modeling of Heteroduplex Formation during PCR from Mixtures of DNA Templates.” *PCR Methods and Applications* 2 (2): 112–16.

Sambrook, Joseph, and David W. Russell. 2006. “Isolation of DNA Fragments from Polyacrylamide Gels by the Crush and Soak Method.” *CSH Protocols* 2006 (1). doi:10.1101/pdb.prot2936.

Sambrook, Joseph, and David William Russell. 2001. *Molecular Cloning: A Laboratory Manual*. CSHL Press.

Sampetrean, Oltea, Shin-ichi Iida, Shinji Makino, Yuriko Matsuzaki, Kikuo Ohno, and Hideyuki Saya. 2009. “Reversible Whole-Organism Cell Cycle Arrest in a Living Vertebrate.” *Cell Cycle* 8 (4): 620–27. doi:10.4161/cc.8.4.7785.

Sánchez-Sánchez, Ana Virginia, Esther Camp, Aránzazu Leal-Tassias, Stuart P. Atkinson, Lyle Armstrong, Manuel Díaz-Llopis, and José L. Mullor. 2010. “Nanog Regulates Primordial Germ Cell Migration Through Cxcr4b.” *Stem Cells* 28 (9): 1457–64. doi:10.1002/stem.469.

Sato, Shin-ichi, Mizuki Watanabe, Yousuke Katsuda, Asako Murata, Dan Ohtan Wang, and Motonari Uesugi. 2015. “Live-Cell Imaging of

Endogenous mRNAs with a Small Molecule." *Angewandte Chemie (International Ed. in English)* 54 (6): 1855–58. doi:10.1002/anie.201410339.

Schuff, Maximilian, Doreen Siegel, Melanie Philipp, Karin Bundschu, Nicole Heymann, Cornelia Donow, and Walter Knöchel. 2012. "Characterization of Danio Rerio Nanog and Functional Comparison to *Xenopus* Vents." *Stem Cells and Development* 21 (8): 1225–38. doi:10.1089/scd.2011.0285.

Seferos, Dwight S., David A. Giljohann, Haley D. Hill, Andrew E. Prigodich, and Chad A. Mirkin. 2007. "Nano-Flares: Probes for Transfection and mRNA Detection in Living Cells." *Journal of the American Chemical Society* 129 (50): 15477–79. doi:10.1021/ja0776529.

Seidel, Claus A. M., Andreas Schulz, and Markus H. M. Sauer. 1996. "Nucleobase-Specific Quenching of Fluorescent Dyes. 1. Nucleobase One-Electron Redox Potentials and Their Correlation with Static and Dynamic Quenching Efficiencies." *The Journal of Physical Chemistry* 100 (13): 5541–53. doi:10.1021/jp951507c.

Semotok, Jennifer L., Hua Luo, Ramona L. Cooperstock, Angelo Karaiskakis, Heli K. Vari, Craig A. Smibert, and Howard D. Lipshitz. 2008. "*Drosophila* Maternal Hsp83 mRNA Destabilization Is Directed by Multiple SMAUG Recognition Elements in the Open Reading Frame." *Molecular and Cellular Biology* 28 (22): 6757–72. doi:10.1128/MCB.00037-08.

Shibata, Yasushi, Bindhu Paul-Prasanth, Aya Suzuki, Takeshi Usami, Masatoshi Nakamoto, Masaru Matsuda, and Yoshitaka Nagahama. 2010. "Expression of Gonadal Soma Derived Factor (GSDF) Is Spatially and Temporally Correlated with Early Testicular Differentiation in Medaka." *Gene Expression Patterns: GEP* 10 (6): 283–89. doi:10.1016/j.gep.2010.06.005.

Shima, Akihiro, and Hiroshi Mitani. 2004. "Medaka as a Research Organism: Past, Present and Future." *Mechanisms of Development*, Medaka, 121 (7–8): 599–604. doi:10.1016/j.mod.2004.03.011.

Singh, Sanjay K., Alexei A. Koshkin, Jesper Wengel, and Poul Nielsen. 1998. "LNA (locked Nucleic Acids): Synthesis and High-Affinity Nucleic Acid Recognition." *Chemical Communications*, no. 4 (January): 455–56. doi:10.1039/A708608C.

Slack, Christine, Anne E. Warner, and R. L. Warren. 1973. "The Distribution of Sodium and Potassium in Amphibian Embryos during Early Development." *The Journal of Physiology* 232 (2): 297–312. doi:10.1113/jphysiol.1973.sp010271.

Socher, Elke, Lucas Bethge, Andrea Knoll, Nadine Jungnick, Andreas Herrmann, and Oliver Seitz. 2008. "Low-Noise Stemless PNA Beacons for Sensitive DNA and RNA Detection." *Angewandte Chemie International Edition* 47 (49): 9555–59. doi:10.1002/anie.200803549.

Sokol, Deborah L., Xiaolin Zhang, Ponzy Lu, and Alan M. Gewirtz. 1998. "Real Time Detection of DNA·RNA Hybridization in Living Cells." *Proceedings of the National Academy of Sciences* 95 (20): 11538–43. doi:10.1073/pnas.95.20.11538.

Song, Wenjiao, Rita L. Strack, Nina Svensen, and Samie R. Jaffrey. 2014. "Plug-and-Play Fluorophores Extend the Spectral Properties of

Spinach." *Journal of the American Chemical Society* 136 (4): 1198–1201. doi:10.1021/ja410819x.

Stein, David, Ernest Foster, Sung-Ben Huang, Dwight Weller, and James Summerton. 1997. "A Specificity Comparison of Four Antisense Types: Morpholino, 2′-O-Methyl RNA, DNA, and Phosphorothioate DNA." *Antisense and Nucleic Acid Drug Development* 7 (3): 151–57. doi:10.1089/oli.1.1997.7.151.

Summerton, James. 1999. "Morpholino Antisense Oligomers: The Case for an RNase H-Independent Structural Type." *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* 1489 (1): 141–58. doi:10.1016/S0167-4781(99)00150-5.

Summerton, James, and Dwight Weller. 1997. "Morpholino Antisense Oligomers: Design, Preparation, and Properties." *Antisense and Nucleic Acid Drug Development* 7 (3): 187–95. doi:10.1089/oli.1.1997.7.187.

Swiger, Roy R., and James D. Tucker. 1996. "Fluorescence in Situ Hybridization: A Brief Review." *Environmental and Molecular Mutagenesis* 27 (4): 245–54. doi:10.1002/(SICI)1098-2280(1996)27:4<245::AID-EM1>3.0.CO;2-C.

Tadros, Wael, and Howard D. Lipshitz. 2009. "The Maternal-to-Zygotic Transition: A Play in Two Acts." *Development* 136 (18): 3033–42. doi:10.1242/dev.033183.

Tanaka, Minoru, Masato Kinoshita, Daisuke Kobayashi, and Yoshitaka Nagahama. 2001. "Establishment of Medaka (*Oryzias latipes*) Transgenic Lines with the Expression of Green Fluorescent Protein Fluorescence Exclusively in Germ Cells: A Useful Model to Monitor Germ Cells in a Live Vertebrate." *Proceedings of the National Academy of Sciences* 98 (5): 2544–49. doi:10.1073/pnas.041315498.

Tanner, Minna, David Gancberg, Angelo Di Leo, Denis Larsimont, Ghizlane Rouas, Martine J. Piccart, and Jorma Isola. 2000. "Chromogenic in Situ Hybridization: A Practical Alternative for Fluorescence in Situ Hybridization to Detect HER-2/neu Oncogene Amplification in Archival Breast Cancer Samples." *The American Journal of Pathology* 157 (5): 1467–72. doi:10.1016/S0002-9440(10)64785-2.

Tay, Yvonne, Jinqiu Zhang, Andrew M. Thomson, Bing Lim, and Isidore Rigoutsos. 2008. "MicroRNAs to Nanog, Oct4 and Sox2 Coding Regions Modulate Embryonic Stem Cell Differentiation." *Nature* 455 (7216): 1124–28. doi:10.1038/nature07299.

Temminck, J.C., and H Schlegel. 1850. "Fauna Japonica." In *Fauna Japonica*, Siebold, Ph.F. von, Pisces tabulae:224–25. A. Arnz et socios, Leiden, the Netherlands.

Tercero, Napoleon, Kang Wang, Ping Gong, and Rastislav Levicky. 2009. "Morpholino Monolayers: Preparation and Label-Free DNA Analysis by Surface Hybridization." *Journal of the American Chemical Society* 131 (13): 4953–61. doi:10.1021/ja810051q.

Thomas, Holly R., Stefanie M. Percival, Bradley K. Yoder, and John M. Parant. 2014. "High-Throughput Genome Editing and Phenotyping Facilitated by High Resolution Melting Curve Analysis." *PLoS ONE* 9 (12): e114632. doi:10.1371/journal.pone.0114632.

Thorball, N. 1981. “FITC-Dextran Tracers in Microcirculatory and Permeability Studies Using Combined Fluorescence Stereo Microscopy, Fluorescence Light Microscopy and Electron Microscopy.” *Histochemistry* 71 (2): 209–33. doi:10.1007/BF00507826.

Triques, Karine, Elodie Piednoir, Marion Dalmais, Julien Schmidt, Christine Le Signor, Mark Sharkey, Michel Caboche, Bénédicte Sturbois, and Abdelhafid Bendahmane. 2008. “Mutation Detection Using ENDO1: Application to Disease Diagnostics in Humans and TILLING and Eco-TILLING in Plants.” *BMC Molecular Biology* 9 (April): 42. doi:10.1186/1471-2199-9-42.

Tsourkas, Andrew, Mark A. Behlke, and Gang Bao. 2002a. “Structure–function Relationships of Shared-stem and Conventional Molecular Beacons.” *Nucleic Acids Research* 30 (19): 4208–15. doi:10.1093/nar/gkf536.

———. 2002b. “Hybridization of 2′-O-Methyl and 2′-Deoxy Molecular Beacons to RNA and DNA Targets.” *Nucleic Acids Research* 30 (23): 5168–74.

Tsourkas, Andrew, Mark A. Behlke, Scott D. Rose, and Gang Bao. 2003. “Hybridization Kinetics and Thermodynamics of Molecular Beacons.” *Nucleic Acids Research* 31 (4): 1319–30. doi:10.1093/nar/gkg212.

Tyagi, Sanjay. 2009. “Imaging Intracellular RNA Distribution and Dynamics in Living Cells.” *Nature Methods* 6 (5): 331–38. doi:10.1038/nmeth.1321.

Tyagi, Sanjay, Diana P. Bratu, and Fred Russell Kramer. 1998. “Multicolor Molecular Beacons for Allele Discrimination.” *Nature Biotechnology* 16 (1): 49–53. doi:10.1038/nbt0198-49.

Tyagi, S., and F. R. Kramer. 1996. “Molecular Beacons: Probes That Fluoresce upon Hybridization.” *Nature Biotechnology* 14 (3): 303–8. doi:10.1038/nbt0396-303.

Uchiyama, Hisatoshi, Ken’ichi Hirano, Masaki Kashiwasake-Jibu, and Kazunari Taira. 1996. “Detection of Undegraded Oligonucleotides in Vivo by Fluorescence Resonance Energy Transfer NUCLEASE ACTIVITIES IN LIVING SEA URCHIN EGGS.” *Journal of Biological Chemistry* 271 (1): 380–84. doi:10.1074/jbc.271.1.380.

Urbanek, Martyna O, Paulina Galka-Marciniak, Marta Olejniczak, and Wlodzimierz J Krzyzosiak. 2014. “RNA Imaging in Living Cells – Methods and Applications.” *RNA Biology* 11 (8): 1083–95. doi:10.4161/rna.35506.

Valencia-Burton, Maria, Ron M. McCullough, Charles R. Cantor, and Natalia E. Broude. 2007. “RNA Visualization in Live Bacterial Cells Using Fluorescent Protein Complementation.” *Nature Methods* 4 (5): 421–27. doi:10.1038/nmeth1023.

Vet, Jacqueline A. M., and Salvatore A. E. Marras. 2005. “Design and Optimization of Molecular Beacon Real-Time Polymerase Chain Reaction Assays.” *Methods in Molecular Biology (Clifton, N.J.)* 288: 273–90.

Vize, Peter D., Kyle E. McCoy, and Xiaolan Zhou. 2009. “Multichannel Wholemount Fluorescent and Fluorescent/chromogenic in Situ Hybridization in *Xenopus* Embryos.” *Nature Protocols* 4 (6): 975–83. doi:10.1038/nprot.2009.69.

Voeltz, Gia K., and Joan A. Steitz. 1998. “AUUUA Sequences Direct mRNA Deadenylation Uncoupled from Decay during *Xenopus* Early Development.” *Molecular and Cellular Biology* 18 (12): 7537–45.

Vouillot, Léna, Aurore Thélie, and Nicolas Pollet. 2015. “Comparison of T7E1 and Surveyor Mismatch Cleavage Assays to Detect Mutations Triggered by Engineered Nucleases.” *G3: Genes|Genomes|Genetics* 5 (3): 407–15. doi:10.1534/g3.114.015834.

Wang, Lin, Chaoyong James Yang, Colin D. Medley, Steven A. Benner, and Weihong Tan. 2005. “Locked Nucleic Acid Molecular Beacons.” *Journal of the American Chemical Society* 127 (45): 15664–65. doi:10.1021/ja052498g.

Wang, Q. Tian, Karolina Piotrowska, Maria Anna Ciemerych, Ljiljana Milenkovic, Matthew P. Scott, Ronald W. Davis, and Magdalena Zernicka-Goetz. 2004. “A Genome-Wide Study of Gene Activity Reveals Developmental Signaling Pathways in the Preimplantation Mouse Embryo.” *Developmental Cell* 6 (1): 133–44. doi:10.1016/S1534-5807(03)00404-0.

Wang, Tiansu, and Yunhan Hong. 2014. “Direct Gene Disruption by TALENs in Medaka Embryos.” *Gene* 543 (1): 28–33. doi:10.1016/j.gene.2014.04.013.

Wei, Zheng, Robert C. Angerer, and Lynne M. Angerer. 2006. “A Database of mRNA Expression Patterns for the Sea Urchin Embryo.” *Developmental Biology* 300 (1): 476–84. doi:10.1016/j.ydbio.2006.08.034.

Wiedenheft, Blake, Samuel H. Sternberg, and Jennifer A. Doudna. 2012. “RNA-Guided Genetic Silencing Systems in Bacteria and Archaea.” *Nature* 482 (7385): 331–38. doi:10.1038/nature10886.

Wiegant, J., C.C. Wiesmeijer, J.M.N. Hoovers, E. Schuuring, A. d’Azzo, J. Vrolijk, H.J. Tanke, and A.K. Raap. 1993. “Multiple and Sensitive Fluorescence in Situ Hybridization with Rhodamine-, Fluorescein-, and Coumarin-Labeled DNAs.” *Cytogenetic and Genome Research* 63 (1): 73–76. doi:10.1159/000133507.

Wittbrodt, Joachim, Akihiro Shima, and Manfred Schartl. 2002. “Medaka — a Model Organism from the Far East.” *Nature Reviews Genetics* 3 (1): 53–64. doi:10.1038/nrg704.

Wolfe, Scot A., Lena Nekludova, and Carl O. Pabo. 2000. “DNA Recognition by Cys2His2 Zinc Finger Proteins.” *Annual Review of Biophysics and Biomolecular Structure* 29 (1): 183–212. doi:10.1146/annurev.biophys.29.1.183.

Wu, Bin, Jiahao Chen, and Robert H. Singer. 2014. “Background Free Imaging of Single mRNAs in Live Cells Using Split Fluorescent Proteins.” *Scientific Reports* 4 (January). doi:10.1038/srep03615.

Wu, Cuichen, Sena Cansiz, Liqin Zhang, I-Ting Teng, Liping Qiu, Juan Li, Yuan Liu, et al. 2015. “A Nonenzymatic Hairpin DNA Cascade Reaction Provides High Signal Gain of mRNA Imaging inside Live Cells.” *Journal of the American Chemical Society* 137 (15): 4900–4903. doi:10.1021/jacs.5b00542.

Wu, Cuichen Sam, Lu Peng, Mingxu You, Da Han, Tao Chen, Kathryn R. Williams, Chaoyong James Yang, and Weihong Tan. 2012. “Engineering Molecular Beacons for Intracellular Imaging.” *International*

Journal of Molecular Imaging 2012 (November): e501579. doi:10.1155/2012/501579.

Xi, Chuanwu, Michal Balberg, Stephen A. Boppart, and Lutgarde Raskin. 2003. “Use of DNA and Peptide Nucleic Acid Molecular Beacons for Detection and Quantification of rRNA in Solution and in Whole Cells.” *Applied and Environmental Microbiology* 69 (9): 5673–78.

Yang, Chaoyong James, Karen Martinez, Hui Lin, and Weihong Tan. 2006. “Hybrid Molecular Probe for Nucleic Acid Analysis in Biological Samples.” *Journal of the American Chemical Society* 128 (31): 9986–87. doi:10.1021/ja0618346.

Yang, Chaoyong James, Lin Wang, Yanrong Wu, Youngmi Kim, Colin D. Medley, Hui Lin, and Weihong Tan. 2007. “Synthesis and Investigation of Deoxyribonucleic Acid/locked Nucleic Acid Chimeric Molecular Beacons.” *Nucleic Acids Research* 35 (12): 4030–41. doi:10.1093/nar/gkm358.

Yang, Hui, Haoyi Wang, Chikdu S. Shivalila, Albert W. Cheng, Linyu Shi, and Rudolf Jaenisch. 2013. “One-Step Generation of Mice Carrying Reporter and Conditional Alleles by CRISPR/Cas-Mediated Genome Engineering.” *Cell* 154 (6): 1370–79. doi:10.1016/j.cell.2013.08.022.

Yeh, H.-Y., M. V. Yates, A. Mulchandani, and W. Chen. 2008. “Visualizing the Dynamics of Viral Replication in Living Cells via Tat Peptide Delivery of Nuclease-Resistant Molecular Beacons.” *Proceedings of the National Academy of Sciences* 105 (45): 17522–25. doi:10.1073/pnas.0807066105.

Yi, Meisheng, Ni Hong, and Yunhan Hong. 2009. “Generation of Medaka Fish Haploid Embryonic Stem Cells.” *Science* 326 (5951): 430–33. doi:10.1126/science.1175151.

Yoon, C., K. Kawakami, and N. Hopkins. 1997. “Zebrafish Vasa Homologue RNA Is Localized to the Cleavage Planes of 2- and 4-Cell-Stage Embryos and Is Expressed in the Primordial Germ Cells.” *Development* 124 (16): 3157–65.

Youil, R, B W Kemper, and R G Cotton. 1995. “Screening for Mutations by Enzyme Mismatch Cleavage with T4 Endonuclease VII.” *Proceedings of the National Academy of Sciences of the United States of America* 92 (1): 87–91.

Yu, Yong A., Kerby Oberg, Gefu Wang, and Aladar A. Szalay. 2003. “Visualization of Molecular and Cellular Events with Green Fluorescent Proteins in Developing Embryos: A Review.” *Luminescence* 18 (1): 1–18. doi:10.1002/bio.701.

Zhang, Y, Z Qu, S Kim, V Shi, B Liao, P Kraft, R Bandaru, Y Wu, L M Greenberger, and I D Horak. 2011. “Down-Modulation of Cancer Targets Using Locked Nucleic Acid (LNA)-Based Antisense Oligonucleotides without Transfection.” *Gene Therapy* 18 (4): 326–33. doi:10.1038/gt.2010.133.

Zhao, Jianxin, Rina Wu, Alfred Au, Abbey Marquez, Yibing Yu, and Zuorong Shi. 2002. “Determination of HER2 Gene Amplification by Chromogenic In Situ Hybridization (CISH) in Archival Breast Carcinoma.” *Modern Pathology* 15 (6): 657–65. doi:10.1038/modpathol.3880582.

Zon, Leonard I., and Randall T. Peterson. 2005. “In Vivo Drug Discovery in the Zebrafish.” *Nature Reviews Drug Discovery* 4 (1): 35–44. doi:10.1038/nrd1606.

Zou, S. 1997. “A Practical Approach to Genetic Screening for Influenza Virus Variants.” *Journal of Clinical Microbiology* 35 (10): 2623–27.

Zu, Yanbing, Aik Leong Ting, Guangshun Yi, and Zhiqiang Gao. 2011. “Sequence-Selective Recognition of Nucleic Acids under Extremely Low Salt Conditions Using Nanoparticle Probes.” *Analytical Chemistry* 83 (11): 4090–94. doi:10.1021/ac2001516.

APPENDIX

Step-by-step procedure of PAGE-HMA method for GE experiments.

1. DNA extraction: Extract and purify genomic DNA from individuals of interest. Avoid cross contamination of different samples.
2. Primers design: Design a pair of primers flanking the gene editing target site to yield PCR products of 200-400 bp.
3. Genomic PCR: For F0 samples, perform PCR with the DNA samples (~50 ng each) using optimal PCR program according to the Tm of primer and the polymerase used; for progenies of the founder fish, refer to step 4. Additionally, one wild-type sample should be included to serve later as a negative control. To reduce artificial mutation, polymerase of higher fidelity is recommended.
4. Heteroduplex formation:
    a) For F0 samples containing disrupted alleles, heteroduplexes are formed naturally during PCR. No additional step is required.
    b) For progenies that might be homozygous, wild-type allele should be added to form heteroduplexes. After PCR, each product should be mixed with wild-type PCR product with a ratio of ~1:1, heated to 94°C for 3 min and slowly returned to room temperature to form heteroduplexes. Alternatively, wild-type genomic DNA sample can be added to the tested sample at 1:1 ratio before PCR, allowing heteroduplexes to form during PCR.
5. PAGE detection:
    a) Prepare a 8% polyacrylamide gel with 1×TBE buffer
    b) Load 4-6 μL PCR product with 6×loading dye to each well of the gel. Run

electrophoresis at 12 V/cm for ~2 h.

c) Stain the gel with 5×GelRed or other DNA gel stains for 20 min and document it on a bio-imaging system.

6. Allelic alteration identification: Compare bands of the sample with those of the negative control. The presence of bands different from the control indicates allelic alterations. If no heteroduplex band can be seen, experiment should be redone with a positive control sample which contains heteroduplex.
7. Repeat step 3 to 6 at least once to rule out false positive signals from PCR artifact.
8. Gel recovery, cloning and sequencing:
   a) Recover DNA of heteroduplexes by cutting bands of heteroduplexes followed by smashing and soaking them in 10-20 μL of 1×Tris-EDTA buffer in 37 °C for overnight (Sambrook and Russell 2006).
   b) Use 1-2 μL of the recovered DNA as templates for one round of grPCR of 30 cycles.
   c) Clone the products into pGEM-T Easy vector. Pick as least 16 colonies for plasmid DNA extraction using the alkaline-SDS mini-prep procedure.
   d) Sequence the insert of the plasmid DNA with proper primer, and compare the result with the wild-type allele.

$$N = \left\lfloor \sqrt{B} + \frac{1}{2} \right\rfloor$$

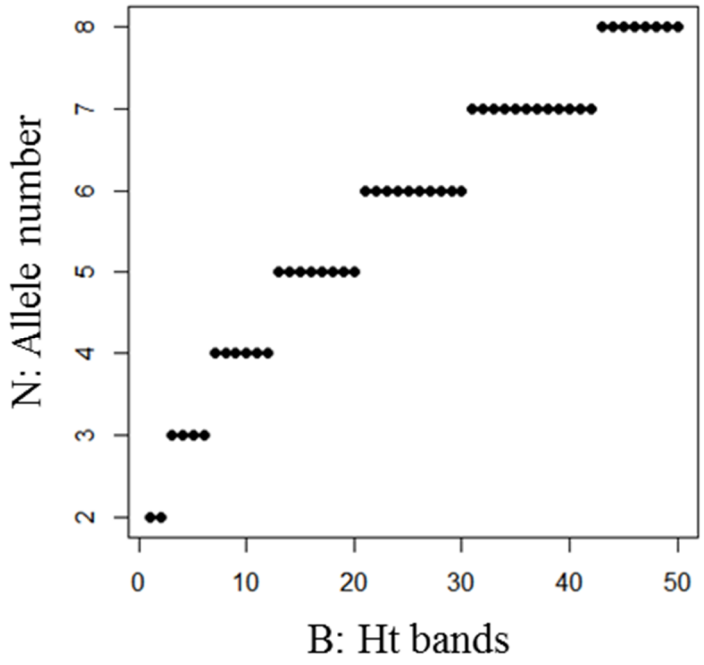


Figure 1. Plot of Equation (3) with Ht band number B from 1 to 50. This plot shows that for each B, Equation (3) returns one corresponding N, which represents the minimum allele number of the mixture in a test.

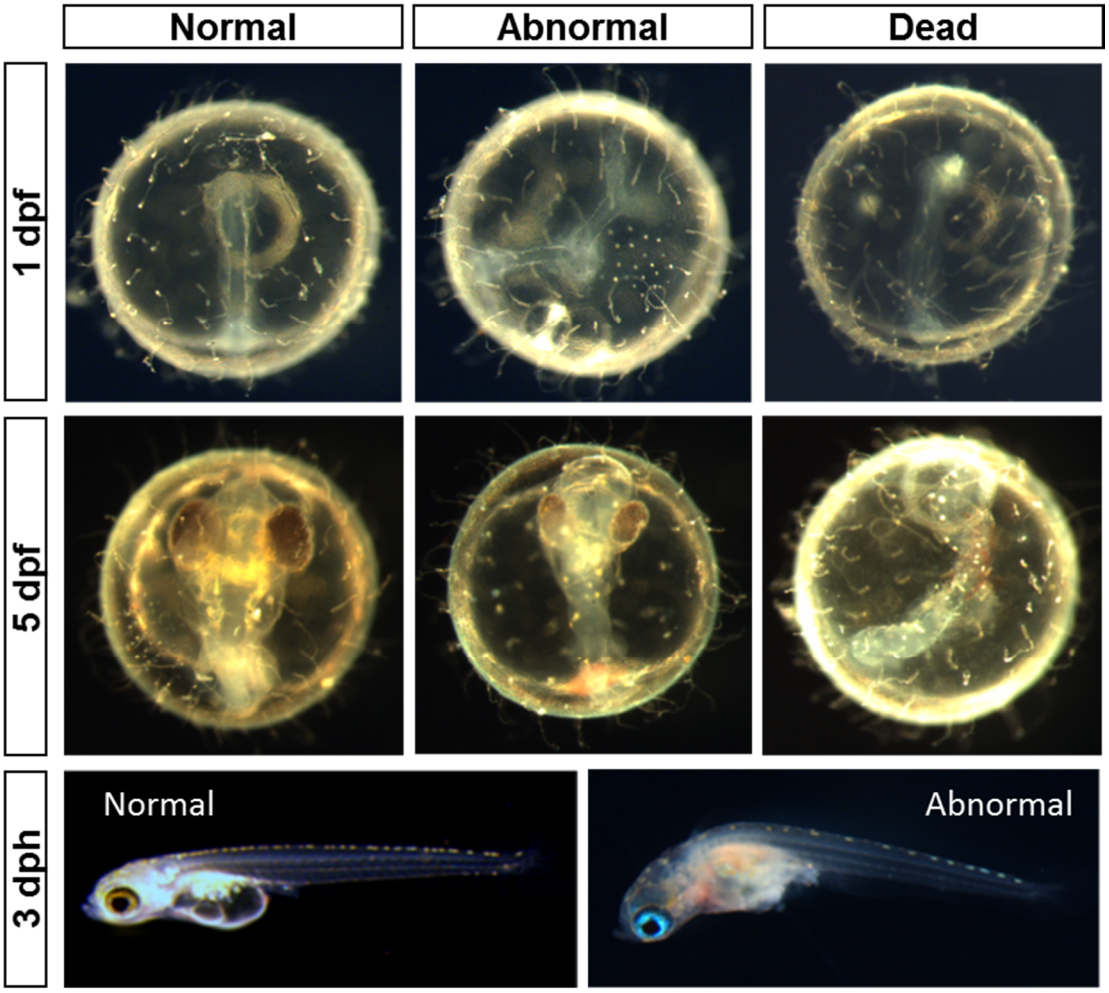


Figure 2. Typical examples of normal, abnormal and dead embryos and fry.

Python script 1. Image processing for stability test.

```
#This script is to calculate the average intensity of green and
red colour in a cropped area of a PNG image file. The green/red
ratio is also calculated.

from PIL import Image
from os import listdir
from statistics import mean,stdev
import numpy as np

#set the input folder as 'direc' of MO and DNA where files of
cropped images locate with name format of:DNA060mR.png,
MO_180mG.png.

def g_r_ratio(direc):
    #averaging the background with BG.png or BR.png, green & red
channel respectively. The two background images are cropped from
the dark background area of the original image.
    bgr = Image.open(direc+'BG.png').convert('L')
    bg = sum(list(bgr.getdata()))/float(bgr.size[0]*bgr.size[1])
    bre = Image.open(direc+'BR.png').convert('L')
    br = sum(list(bre.getdata()))/float(bre.size[0]*bre.size[1])

    #array avr[] store 7 value of each image: the time, green and
red value with stdev (standard deviation), and ratio with stdev,
as denoted in 'header'.
    avr = [[] for i in range(7)]
    filename = [f for f in listdir(direc) if f[-3:]=='png' and
f[0]!='B']
    #g & r store all pixel intensity of green or red
    g = []
    r = []
    for f in filename:
        im = Image.open(direc+f).convert('L')
        d = list(im.getdata())
        if f[-5] == 'G':
            avr[0].append(f[3:-6])
            avr[1].append(mean(d)-bg)
            avr[2].append(stdev([x-bg for x in d]))
            g.append([x-bg for x in d])
        else:
            avr[3].append((mean(d)-br)*0.4)
            avr[4].append(stdev([x-br for x in d]))
            r.append([x-br for x in d])
```

```
    #print the result for checking
    for i in range(len(avr[0])):
        ratio = avr[1][i]/avr[3][i]
        avr[5].append(ratio)
        print(avr[0][i]+' '+str(avr[1][i])+' '+ str(avr[3][i]) + '
'+ str(ratio)[:4])
    #calculate the stdev or each g/r ratio by pixel
    for i in range(len(g)):
        rt = [g[i][j]/float(r[i][j]) for j in range(len(g[i]))]
```

```
        avr[6].append(stdev(rt))
```

```
    #output to file
    header = 'Time\t Green Intensity\t Green Stdev\t Red
Intensity\t Red Stdev\t Green/Red ratio\t Ratio Stdev'
    with open(direc + 'output_data.txt','w') as out:
        out.write(header)
        for i in range(len(avr[0])):
            out.write('\n')
            for j in range(len(avr)):
                out.write(str(avr[j][i])+'\t')
```